PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :		(11) International Publication Number: WO 00/63381
C12N 15/12, C07K 14/47, 16/18, C12N 15/11, G01N 33/68, A61K 38/17, 39/395, 48/00	A1	(43) International Publication Date: 26 October 2000 (26.10.00)

(21) International Application Number: PCT/US00/09488

(22) International Filing Date: 11 April 2000 (11.04.00)

(30) Priority Data:
60/129,553 16 April 1999 (16.04.99) US

(71) Applicant *(for all designated States except US)*: SCIOS, INC. [US/US]; 820 West Maude Avenue, Sunnyvale, CA 94086 (US).

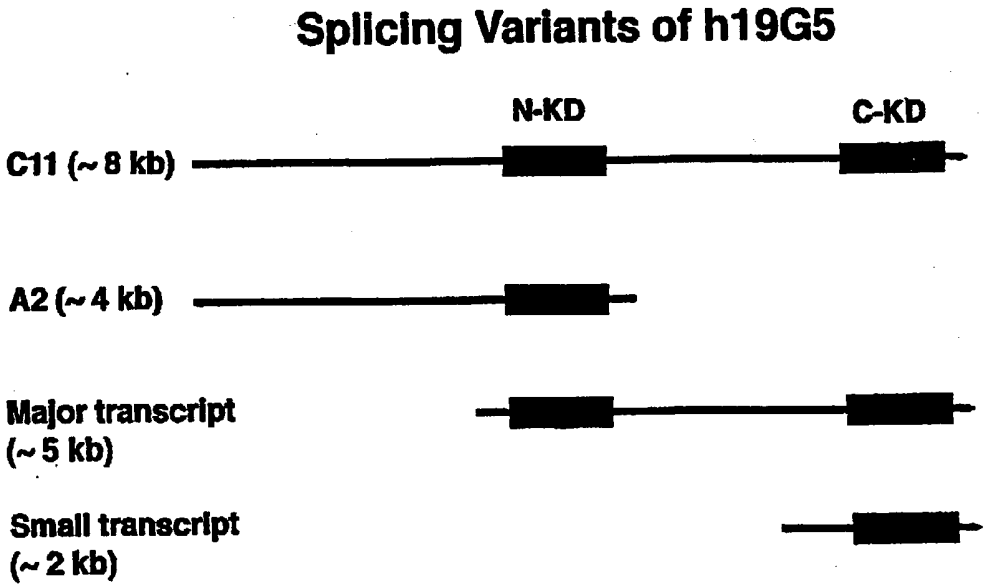
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: ZENG, Wenlin [CN/US]; 2210 Rock Street #24, Mountain View, CA 94043 (US). STANTON, Lawrence [US/US]; 73 Turnsworth Avenue, Redwood City, CA 94062 (US). KONG, Haiyan [CN/US]; Apartment #6, 914 Tamarack Lane, Sunnyvale, CA 94086 (US).

(74) Agent: ALTMAN, Daniel, E.; Knobbe, Martens, Olson and Bear, LLP, 16th Floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), DM, DZ, EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: MAMMALIAN PROTEIN WITH PUTATIVE FUNCTION IN SIGNAL TRANSDUCTION



(57) Abstract

Polypeptides capable of regulating signal transduction, which preferably exhibit kinase activity, or antibodies against such polypeptides that inhibit the interaction of these polypeptides with other mediators of signal transduction, may be used in the identification, prevention or treatment of disease, preferably cardiac disease, in mammalian hosts. In addition, these polypeptides can facilitate the identification or isolation of additional mediators of signal transduction associated with disease, preferably cardiac disease, which in turn may also be used in the identification, prevention or treatment of disease, preferably cardiac disease, in mammals.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

MAMMALIAN PROTEIN WITH PUTATIVE FUNCTION IN SIGNAL TRANSDUCTION

FIELD OF THE INVENTION

The present invention relates to compounds and methods for the identification, prevention or treatment of disease, preferably cardiac disease, in a mammal through the administration of polypeptides capable of regulating signal transduction, which preferably exhibit kinase activity, or antibodies against such polypeptides that inhibit the interaction of these polypeptides with other mediators of signal transduction. In addition, the compounds and methods of the present invention can facilitate the identification or isolation of additional mediators of signal transduction associated with disease, preferably cardiac disease, which in turn may also be used in the identification, prevention or treatment of disease, preferably cardiac disease, in mammals.

BACKGROUND OF THE INVENTION

Certain biological functions, such as growth and differentiation, are tightly regulated by signal transduction pathways within cells. Signal transduction pathways maintain the balanced steady state functioning of a cell. Disease states can arise when signal transduction in a cell breaks down, thereby removing the control that typically exists over cellular functions. Because signal transduction networks regulate a multitude of cellular functions depending upon the cell type, a wide variety of diseases can result from abnormalities in such networks. Devastating diseases such as cancer, autoimmune diseases, allergic reactions, inflammation, neurological disorders and hormone-related diseases can result from abnormal signal transduction. For example, tumors may develop when regulation of cell growth is disrupted.

Despite a long-felt need to understand and discover methods for regulating cells involved in various disease states, the complexity of signal transduction pathways has precluded the development of products and processes for regulating cellular function by manipulating signal transduction pathways in a cell. As such, there remains a need for products and processes that permit the implementation of predictable controls of signal transduction in cells, thus enabling the treatment of various diseases that are caused by abnormal cellular function.

Such diseases may include cardiac diseases, which may include congestive heart failure (CHF), dilated congestive cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve disease, aortic valve disease, tricuspid valve disease, angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary hypertension, arterial hypertension, renovascular hypertension, arteriosclerosis, atherosclerosis, and cardiac tumors. By way of example, CHF is a major cardiac disease associated with extensive morbidity and mortality. Approximately five million individuals in the United States suffer from some form of CHF. Traditionally, treatment of CHF occurs by a series of agents including diuretics, vasodilators, angiotensin converting enzyme inhibitors, ß-adrenergic antagonists, and positive inotropes like digoxin. These drugs, however, principally provide symptomatic relief and typically only extend the life of one suffering from the disease for periods ranging from 6-12 months.

The pathophysiology of CHF is rather complex. Generally, the central hallmark of the disease is the inability of the heart to pump sufficient oxygenated blood to meet the demands of peripheral tissues. Numerous etiologies contribute to the development of CHF, including primary diseases of, or insults to, the myocardium itself, cardiac defects, hypertension, inflammation, kidney disease and vascular disease. These conditions lead to the hypertrophy and remodeling of the cardiac ventricles which, if unchecked, ultimately reduce the mechanical performance of the heart. Forces associated with the inability of the heart to pump blood ultimately lead to the release of neurohormones like catecholamines, renin-angiotensin, aldosterone, endothelin and related factors into the circulation. Elevations in plasma levels of many of these circulating neurohormones have a deleterious impact on the outcome of patients with CHF. Local production of these neurohormonal factors in the heart is believed to contribute centrally to the disease. Thus, an important therapeutic strategy has been to block this neurohormonal axis contributing to the pathogenesis of this disease.

Factors known to contribute centrally to the pathophysiology of heart disease are biosynthesized in the heart itself. These factors are produced in cardiac myocytes, fibroblasts, smooth muscle and endothelial cells, and inflammatory cells associated with the myocardium. For example, the heart contains its own renin-angiotensin system. Blockade of the cardiac renin-angiotensin system may contribute significantly to the therapeutic efficacy of the therapeutic class of agents known as angiotensin converting enzyme (ACE) inhibitors.

The heart also produces other factors including endothelins, bradykinin, adrenomedullin, tumor necrosis factor, transforming growth factors, and natriuretic peptides. Unfortunately, therapeutic strategies are limited to the modulation of such substances, which are already known to contribute to the disease. Indeed, the functional contributions of only a minor fraction of all known secreted factors encoded by the human genome have apparently been defined.

The foregoing shows a need for methods and products involving the prevention or treatment of disease in mammals involving the mediation of signal transduction. The administration of polypeptides capable of regulating signal transduction, which preferably exhibit kinase activity, or antibodies against such polypeptides that inhibit the interaction of these polypeptides with other mediators of signal transduction, in addition to the identification or isolation of additional mediators of signal transduction associated with disease, preferably cardiac disease, which in turn may also be used in the identification, prevention or treatment of disease, preferably cardiac disease, in mammals, can facilitate such prevention or treatment.

SUMMARY OF THE INVENTION

An objective of the present invention is therefore the prevention or treatment of disease, preferably cardiac disease, in mammals through the administration of polypeptides capable of regulating signal transduction, which preferably exhibit kinase activity, or antibodies against such polypeptides that inhibit the interaction of these polypeptides with other mediators of signal transduction, in addition to the identification or isolation of additional

mediators of signal transduction associated with disease, preferably cardiac disease, which in turn may also be used in the identification, prevention or treatment of disease, preferably cardiac disease, in mammals.

In accomplishing these and other objectives, the present invention preferably provides a purified polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction. In a preferred embodiment, the polypeptide is capable of catalyzing the transfer of a phosphate group from a donor molecule to an acceptor molecule.

In another embodiment, the present invention preferably provides an isolated DNA molecule encoding a purified polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction. The present invention may also preferably be an isolated DNA molecule comprising the nucleotide sequence of SEQ ID NOS: 2, 3 or 5.

In yet another embodiment, the present invention preferably provides a vector comprising a DNA molecule encoding a purified polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction. In another aspect, the present invention provides a host cell transformed with such a vector. In one other embodiment, the present invention may preferably provide the above-described transformed host cell, where the host cell produces a polypeptide capable of regulating signal transduction. In a preferred embodiment, the above-described transformed host cell produces a polypeptide capable of catalyzing the transfer of a phosphate group from a donor molecule to an acceptor molecule.

In another preferred embodiment, the present invention provides an isolated antibody against a polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction. In a preferred embodiment, the antibody is a monoclonal antibody. In another preferred embodiment, the antibody is capable of inhibiting the regulation of signal transduction. In yet another preferred embodiment, the antibody is capable of inhibiting the transfer of a phosphate group from a donor molecule to an acceptor molecule.

The present invention may also preferably provide an isolated nucleic acid capable of hybridizing under high stringency conditions to a DNA molecule comprising the nucleotide sequence of SEQ ID NOS: 2, 3 or 5. In a preferred embodiment, this isolated nucleic acid is capable of inhibiting the regulation of signal transduction. In yet another preferred embodiment, this isolated nucleic acid is capable of inhibiting said transfer of said phosphate group from said donor molecule to said acceptor molecule.

In a preferred embodiment, the present invention provides a method of preventing or treating disease in a mammal comprising administering to said mammal an effective amount of material, selected from the group consisting of the polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction, and the antibody against this polypeptide, in a pharmaceutically acceptable sterile vehicle. In a preferred embodiment, the mammal may be a human. In another, the disease may be cardiac disease.

The present invention may also preferably provide a vaccine for preventing disease in a mammal comprising administering to said mammal an effective amount of material, selected from the group consisting of the polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction, and the antibody against this polypeptide, in a pharmaceutically acceptable sterile vehicle. In a preferred embodiment, the mammal may be a human. In another, the disease may be cardiac disease.

In a preferred embodiment, the present invention provides a method of preventing or treating disease in a mammal comprising administering to said mammal syngeneic cells transformed with a vector comprising a DNA molecule encoding a purified polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction, and wherein the transformed syngeneic cells produce a polypeptide capable of regulating signal transduction. In a preferred embodiment, the mammal may be a human. In another, the disease may be cardiac disease.

The present invention may also preferably provide a method of preventing or treating disease in a mammal comprising administering to said mammal syngeneic cells transformed with a vector comprising a DNA molecule encoding a purified polypeptide comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction, and wherein the transformed syngeneic cells produce a polypeptide capable of catalyzing the transfer of a phosphate group from a donor molecule to an acceptor molecule. In a preferred embodiment, the mammal may be a human. In another, the disease may be cardiac disease.

In a preferred embodiment, the present invention may provide a kit for detecting the expression of a protein capable of regulating signal transduction, comprising a polypeptide, which comprises the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction. In a preferred embodiment, this kit further comprises a detectable label selected from the group consisting of colorimetric, enzymatic, chemiluminescent, fluorescent and radioactive labels.

In another preferred embodiment, the present invention may provide a kit for detecting the expression of a protein capable of acting as a donor molecule or an acceptor molecule of a phosphate group comprising a polypeptide, which comprises the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction. In a preferred embodiment, this kit further comprises a detectable label selected from the group consisting of colorimetric, enzymatic, chemiluminescent, fluorescent and radioactive labels.

The present invention may also preferably provide a method for detecting the expression of a protein capable of regulating signal transduction, comprising contacting a sample with a polypeptide, which comprises the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction, and detecting any effect of the sample on an indicator of signal transduction. In a preferred embodiment, the polypeptide is immobilized to a solid support. In another preferred embodiment, the phosphate group is detectably labeled.

In another preferred embodiment, the present invention may provide a method for detecting the expression of a protein capable of acting as a donor molecule or an acceptor molecule of a phosphate group, comprising contacting a

sample with a polypeptide, which comprises the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 wherein the polypeptide is capable of regulating signal transduction, and detecting any effect of the sample on an indicator of signal transduction, and detecting any transfer of the phosphate group. In a preferred embodiment, the polypeptide is immobilized to a solid support. In another preferred embodiment, the phosphate group is detectably labeled.

Other objectives, features, and advantages of the present invention will become apparent from the following detailed description. The detailed description and the specific examples, while indicating preferred embodiments of the invention, are provided by way of illustration only. Accordingly, the present invention also includes those various changes and modifications within the spirit and scope of the invention that may become apparent to those skilled in the art from this detailed description.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 illustrates eight overlapping human cDNA clones of 19G5.

Figure 2 depicts homology alignment of amino acid sequences of human, rat and mouse clones corresponding to 19G5.

Figure 3 depicts the likely gene structure of human 19G5 (H19G5), which reveals at least 10 exons and 9 introns.

Figure 4 is a schematic diagram of four cDNA clones corresponding to splicing variants of H19G5. The longest clone (C11) contains two kinase domains, the N-terminal and the C-terminal kinase domains (N-KD and C-KD).

Figure 5 is a schematic diagram comparing the domain structure of 19G5 and Trio proteins.

Figure 6 is the sequence comparison of the kinase domains of H19G5 (N-terminal and C-terminal kinase domains), Trio, and smooth muscle myosin light chain (SM MLCK).

Figure 7 is the sequence comparison of the guanine nucleotide exchange factor (GEF) domains of H19G5 and Trio.

Figure 8 depicts a Northern blot analysis of various rat tissue samples using rat 19G5 DNA as a probe. Heart tissue reveals two transcripts (2 kb and a 4.4 kb) of R19G5. Multiple bands of weak hybridization are seen in skeletal muscle. Parallel Northern analysis of various rat tissue samples using a β-actin probe suggests approximate equal input of mRNA in various samples and integrity of mRNA.

Figure 9 depicts a Northern blot analysis of various human tissue samples using H19G5 DNA probe. Heart tissue reveals multiple transcripts (5 kb, 3 kb, 2.4 kb and 1.8 kb) derived from H19G5 gene. Skeletal muscle and brain tissues show multiple bands of weak hybridization.

Figure 10 depicts a Southern blot analysis of human genomic DNA using H19G5 DNA as a probe (SEQ ID NO: 2).

Figure 11 shows the subcellular localization of 19G5-GFP fusion proteins in mouse myoblast cell line C2C12. Three 19G5-GFP fusion protein expression constructs were made using three different 19G5 cDNA clones, the longest

human 19G5 clone C11 [h19G5(C11)-GFP], a 2.7 kb clone of human 19G5 containing the C-terminal kinase domain [h19G59F1)-GFP], and the rat 19G5 small transcript [r19G5(S)-GFP]. The control GFP vector and the 19G5-GFP fusion expression constructs were transfected into C2C12 cells. The 19G5-GFP fusion proteins' localization was detected using confocal microscopy.

Figure 12 shows that H19G5 protein expressed by the clone C11 binds to the small G protein Cdc42. The lysate of 293 EBNA cells transfected with H19G5-C11 expression construct was incubated with GST-Cdc42 immobilized on glutathione-agarose. After washing, the complex was resuspended in SDS sample buffer, boiled and run on a SDS-PAGE, and Western blotted with an anti-H19G5 monoclonal antibody.

Figure 13 shows phase contrast micrographs of C2C12 myoblasts undergoing differentiation into myotubes. Undifferentiated C2C12 cells are shown in Day 0. Differentiation of C2C12, induced by placing in 2% horse serum, is apparent on Day 3 and Day 5 as shown.

Figure 14 shows induction of myogenin protein expression when C2C12 myoblast cells are induced to differentiate into myotubes. Myogenin is a marker of myotubes.

Figure 15 shows induction of 19G5 RNA expression during the differentiation of myoblasts into myotubes.

Figure 16 shows inhibition of the induction of 19G5 expression by TFG-β during C2C12 differentiation.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Those skilled in the art will recognize that the products and methods embodied in the present invention may be applied to a variety of systems, constructed with various materials using various methods. Accordingly, the present invention is not limited to any particular environment, and the following description of specific embodiments of the present invention are for illustrative purposes only.

The present invention preferably provides methods for the prevention or treatment of disease, preferably cardiac disease, in mammals through the administration of polypeptides capable of regulating signal transduction, which preferably exhibit kinase activity, or antibodies against such polypeptides that inhibit the interaction of these polypeptides with other mediators of signal transduction, in addition to the identification or isolation of additional mediators of signal transduction associated with disease, preferably cardiac disease, which in turn may also be used in the identification, prevention or treatment of disease, preferably cardiac disease, in mammals. The cardiac diseases according to the present invention may include congestive heart failure (CHF), dilated congestive cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve disease, aortic valve disease, tricuspid valve disease, angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary hypertension, arterial hypertension, renovascular hypertension, arteriosclerosis, atherosclerosis, and cardiac tumors.

An embodiment of the invention is a purified polypeptide comprising an amino acid sequence having at least 95% sequence identity with the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9. As used herein, polypeptide refers to a linear series of amino acid residues connected to one another by peptide bonds between the alpha-amino

groups and carboxy groups of adjacent amino acid residues. Additional covalent bonds between portions of the peptide may also be present to restrain the conformation of the molecule, such as amide and disulfide bonds. When used herein, polypeptide also refers to a linear series of amino acid residues connected one to the other as in a peptide. The term synthetic peptide means a chemically derived chain of amino acid residues linked together by peptide bonds that is free of naturally occurring proteins and fragments thereof.

The one and three-letter symbols used to represent the amino acid residues in the polypeptides of the present invention are those symbols commonly used in the art. The amino acid residues are preferred to be in the L isomeric form. However, residues in the D isomeric form may be substituted for any L-amino acid, as long as the desired functional property of signal transduction mediation is retained by the peptide. The one and three-letter symbols used herein refer to the following amino acids: Ser (S) is serine; Ile (I) is isoleucine; Gln (Q) is glutamine; Phe (F) is phenylalanine; His (H) is histidine; Trp (W) is tryptophan; Lys (K) is lysine; Asn (N) is asparagine; Leu (L) is leucine; Gly (G) is glycine; Thr (T) is threonine; Asp (D) is aspartic acid; Arg (R)is arginine; and Ala (A) is alanine.

Polypeptides of the present invention include variants, fragments and chemical derivatives of the polypeptides comprising the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9 as long as they are capable of mediating signal transduction. Polypeptides thus may include soluble peptides, Ig-tailed fusion peptides (including immunoadhesions), members of random peptide libraries (*see, e.g.*, Lam, K.S. *et al.*, *Nature* 354:82-84 (1991); Houghten, R. *et al.*, *Nature* 354:84-86 (1991)), combinatorial chemistry-derived molecular libraries made of D-and/or L-configuration amino acids, and phosphopeptides (including members of random or partially degenerate, directed phosphopeptide libraries, *see, e.g.*, Songyang, Z. *et al.*, *Cell* 72:767-778(1993)).

Polypeptides of the present invention may also include polypeptides that can be isolated from nature or can be produced by recombinant and/or synthetic means. Such native sequence polypeptides specifically refers to naturally-occurring truncated or secreted forms (*e.g.*, an extracellular domain sequence), as well as naturally occurring variant forms (*e.g.*, alternatively spliced forms), and naturally occurring allelic variants of the named polypeptides.

The term variant refers to any polypeptide having an amino acid sequence, in comparison to the amino acid sequences of the polypeptides of the present invention, in which one or more amino acids have been substituted with other amino acids; where the substituted amino acids allow or require the polypeptide to assume the equilibrium conformation of the domain of the parent protein. Often, cysteine, lysine and glutamic acid will be used for their side chains which can form covalent linkages to restrict the conformation of a peptide. The term variant refers to any polypeptide in which one or more amino acids are added and/or substituted and/or deleted and/or inserted at the N- or C-terminus or anywhere within the corresponding native sequence, and which retains signal transduction mediation activity of the corresponding native polypeptide. The variants herein preferably comprise a sequence that has at least about 80% sequence identity, more preferably at least about 85% sequence identity, even more preferably at least about 90% sequence identity, most preferably at least about 95% sequence identity, with the amino acid sequence of SEQ ID NOS: 1, 4, 6, 7, 8 or 9.

In such amino acid sequences, one or more amino acids in the fundamental sequence may preferably be substituted with another amino acid(s), the charge and polarity of which is similar to that of the native amino acid, *i.e.*, a conservative amino acid substitution, resulting in a silent change. Substitutes for an amino acid within the fundamental polypeptide sequence can be selected from other members of the class to which the naturally occurring amino acid belongs. Amino acids can be divided into the following four groups: (1) acidic amino acids; (2) basic amino acids; (3) neutral polar amino acids; and (4) neutral non-polar amino acids. Representative amino acids within these various groups include, but are not limited to: (1) acidic (negatively charged) amino acids such as aspartic acid and glutamic acid; (2) basic (positively charged) amino acids such as arginine, histidine, and lysine; (3) neutral polar amino acids such as glycine, serine, threonine, cyteine, cystine, tyrosine, asparagine, and glutamine; (4) neutral nonpolar (hydrophobic) amino acids such as alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine.

The term variant shall also include any polypeptide having one or more amino acids deleted from or added to an amino acid sequence of a mediator of signal transduction, but which still retains signal transduction mediation activity. The term fragment shall refer to any shorter version of the polypeptides herein, wherein the fragment is capable of mediating signal transduction.

Sequence identity is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in a native polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. The % sequence identity values are preferably generated by the NCBI BLAST2.0 software as defined by Altschul et al., (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.*, 25:3389-3402. The parameters are set to default values, with the exception of the Penalty for mismatch, which is set to -1. Other algorithms, such as GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package (Genetics Computer Groups, 575 Science Dr., Madison, WI), are also suitable. The selection of the non-default parameters to achieve maximum sequence identity is well within the skill of a person skilled in the art.

Antibodies of the present invention may include any polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, $F(ab')_2$ and FAb fragments, and epitope-binding fragments thereof.

Without further elaboration, one skilled in the art with the preceding description can utilize the present invention to its fullest extent. The following examples are illustrative only, and not intended to limit the remainder of the disclosure in any way.

Example 1 Isolation and characterization of 19G5 clones from rat, mouse and human

Isolation of partial and full-length 19G5 clones from rat, mouse and human

Microarray technology was used to identify genes that are differentially expressed in normal and diseased rat heart. A sequence designated as 19G5 was down regulated in 12-week myocardial infarct (MI) rat hearts. A rat cDNA clone corresponding to 19G5 (R19G5) was isolated and nucleotide sequence determined. The deduced amino acid sequence of the clone revealed homology to the catalytic domain of kinases, thus suggesting that the protein product of R19G5 might be involved in signal transduction. A hybridization of multiple tissue Northern blot with the R19G5 probe showed that the gene is highly expressed in heart as a 2 kb and 4.4 kb transcripts (Figure 8).

A full length cDNA for the 2 kb R19G5 transcript was cloned using 5' RACE technique. The R19G5 has an open reading frame of 1644 base pairs which encodes a protein of 548 amino acids (SEQ ID NO: 8)

A mouse 19G5 (M19G5) EST clone was identified by searching the EST database. Northern blot hybridization using the M19G5 EST clone as probe detected a major transcript of 2 kb in heart. There is also a low level of expression in lung. The smeary hybridization was also detected as in the rat and human but was less prominent. M19G5 was also expressed in 17-day old mouse embryos suggesting that it may play a role in embryo development. Sequence analysis of the M19G5 clone showed that it is 1900 base pairs long and has an open reading frame of 1644 base pairs coding for a polypeptide of 548 amino acids (SEQ ID NO: 9), suggesting that it likely represents full length 2 kb transcript.

A number of human cDNA clones corresponding to 19G5 (H19G5) were isolated from human cDNA library using R19G5 as a probe. Figure 1 shows eight overlapping cDNA clones of H19G5. Extensive overlap among these clones helped build a consensus nucleotide sequence (SEQ ID NO: 2) designated as H19G5 contig. The sequence corresponds to a major transcript (~5 kb) expressed in human heart. The sequence of the contig revealed that it is complete at the 3'-end since it contains a polyadenylation signal (AATAAA) as well as polyA residues at the 3'-end. It has a potential open reading frame, coding for 1351 amino acid residues, that extends beyond the 5'-end of the contig indicating that it is incomplete at the 5'-end. The deduced amino acid sequence (SEQ ID NO: 1) revealed a protein kinase domain at the C-terminal end (amino acid residues 1056 to 1309) and also a partial protein kinase domain towards the N-terminal end (amino acid residues 1 to 105) of this truncated clone. The sequence information of this contig (SEQ ID NO: 2) was used to devise antisense primers corresponding to the 5'-end, which were used in 5' RACE (rapid amplification of cDNA ends) to isolate cDNA clones with longer inserts. A cDNA clone containing full-length coding sequence was isolated. The nucleotide sequence of the clone (SEQ ID NO: 3) revealed an open reading frame that could potentially code for 1667 amino acids long full-length polypeptide (SEQ ID NO: 4).

Using the 5' RACE technique, the full coding region for one of the larger H19G5 (3 kb or 5 kb) transcripts was cloned. It has an open reading frame of 2613 base pairs encoding a protein of 811 amino acids.

Figure 2 shows the amino acid sequence alignment of the three full-length 19G5 proteins. Both the R19G5 (SEQ ID NO: 8) and M19G5 (SEQ ID NO: 9) proteins overlap with the C-terminal two-third of a splicing variant of the H19G5 protein (SEQ ID NO: 7). The identity between the R19G5 and M19G5 proteins is 97%. The three proteins are 85% identical in their sequences.

In order to understand the genomic organization of H19G5 gene, phage genomic library was screened and eleven H19G5 genomic clones were isolated. Restriction enzyme mapping of these clones detected no polymorphism, suggesting that H19G5 may be a single copy gene. The entire genomic DNA region encompassing H19G5 was sequenced and found to contain at least 10 exons and 9 introns as shown in Figure 3.

One of the H19G5 cDNA clones contained deletion of two amino acid residues (alanine and proline) in the kinase domain. Sequence analysis of the genomic H19G5 DNA revealed an intron located immediately upstream of the alanine residue. There are two adjacent splicing acceptor sites that are four nucleotides apart at the 3' end of the intron. A splicing event using the first acceptor site generates a protein that includes the two residues, alanine and proline. Utilization of the second splicing acceptor site creates a H19G5 protein with the two amino acids deleted in its kinase domain. These two residues are highly conserved among all kinases. This deletion form of H19G5 protein may thus exhibit reduced or no kinase activity at all, and may serve as a dominant negative inhibitor of the kinase activity of H19G5. This provides a possible mechanism to regulate the activity of H19G5 *in vivo*.

Isolation of cDNA clones of splicing variants of H19G5

Multiple transcripts were detected in Northern analysis of human and rat cells, particularly in heart tissues (As shown in Figs. 8 and 9). However, the restriction analysis of the genomic clones and Southern analysis of genomic DNA suggested single-copy nature of H19G5 gene. These observations indicated the possibility of alternative splicing as a source of multiple transcripts. Four cDNA clones representing various splicing variants were isolated and are schematically shown in Figure 4. Complete nucleotide sequence (SEQ ID NO: 5) of the longest clone C11 was determined and the amino acid sequence of a large open reading frame contained therein was deduced (SEQ ID NO: 6). The large ORF has a potential to encode a protein of 2596 amino acid residues. The sequence analysis revealed the presence of a number of structurally and functionally important domains in H19G5. The presence of some of the domains strengthens the possibility of the involvement of H19G5 in signal transduction. For example, H19G5 has two kinase domains, one at the N-terminal (1094 to 1351 amino acid residues, N-KD) and the other at the C-terminal (2301 to 2553 amino acid residues, C-KD) end. H19G5 also has a Guanine nucleotide Exchange Factor (GEF) domain (325 to 504 amino acid residues), and a Pleckstrin Homology (PH) domain (532 to 634 amino acid residues). Additionally, H19G5 contains five Immunoglobulin (Ig) like domains distributed throughout the sequence.

Structural and functional features of full-length H19G5 protein sequence

The full-length H19G5 amino acid sequence (SEQ ID NO: 6) shows sequence similarity to some functionally important domains of a protein called Trio. The comparison of the domain structure of H19G5 and Trio is schematically depicted in Figure 5. Both kinase domains of H19G5 (N-KD and C-KD) are homologous to a single kinase domain of Trio as well as to a kinase domain of smooth muscle myosin light chain kinase (SM MLCK) as shown in Figure 6. The identity between H19G5 N-KD and Trio KD is about 40% and between H19G5 N-KD and SM MLCK is

about 38%. The sequence identity between the two kinase domains of H19G5 is about 30%. A single GEF domain of H19G5 is homologous to both GEF domains (GEF-D1 and GEF-D2) of Trio as shown in Figure 7.

Proteins containing GEF domains are involved in signal transduction (for a review, see Cherfils and Chardin, *Trends Biochem. Sci.* 24: 306-311 [1999]). GEF domains promote exchange of GTP for GDP on GTP-binding proteins (G proteins) and thereby positively regulate their activities. As these proteins harbor intrinsic GTPase activity, they are also referred to as GTPases. These small G proteins, as opposed to trimeric G proteins, belong to a superfamily of Ras-like proteins. These proteins are bound to the inner face of plasma membrane, and usually exist in GDP-bound "inactive" state. When a ligand interacts with a membrane bound receptor, alteration of conformation allows the receptor to interact with a G protein. This interaction results in conformational change in the G protein that weakens the affinity for GDP and leads to replacement of GDP with GTP. This nucleotide exchange is greatly accelerated or promoted by proteins containing GEF domains. Once bound to GTP, the G proteins assume an "active" state in which they interact with the downstream effectors and facilitate transduction of signal from membrane to the nucleus. However, the activity of the G proteins is tightly controlled as their intrinsic GTPase activity rapidly hydrolyzes bound GTP into GDP and restores "inactive" status. Thus, G proteins function as molecular switches in signal transduction. A number of membrane receptors operate through G proteins. The downstream effectors of the activated G proteins include various protein kinases constituting a cascade of protein phosphorylation that brings about a desired change in gene expression.

As discussed in the preceding section, GEF domains play a critical role in signal transduction by controlling the activation of G proteins. Trio is a complex protein possessing two GEF domains, each with adjacent pleckstrin homology (PH) domains and Src Homology-3 (SH3) domains, a protein serine/threonine kinase domain with an adjacent immunoglobulin-like domain and multiple spectrin-like domains (Medley *et al., J. Cell Sci.* 112: 1825-1834 [1999]). Trio cDNA clone was isolated by virtue of its ability to interact with protein tyrosine phosphatase (PTP) domain of a protein called LAR (Debant *et al., Proc. Natl. Acad. Sci. USA* 93: 5466-5471 [1996]). LAR is a broadly expressed transmembrane protein tyrosine phosphatase comprised of a cell adhesion-like extracellular region and two intracellular PTPase domains, and is proposed to regulate cell-matrix interactions (Mourey and Dixon, *Curr. Opin. Gen. Dev.* 4: 31-39 [1994]). Trio represents a unique member of the Rho-GEFs family possessing two functional GEF domains of distinct specificities. For example, GEF1 is specifically active on Rac1 GTPase, while GEF2 targets RhoA GTPase (Debant *et al, supra*). This unique feature allows Trio to link Rho and Rac specific signalling pathways *in vivo*.

The Rho family of Ras-like GTPases includes Rac (1, 2 and 3), RhoG, Cdc42Hs, TC10, TTF/RhoH, Rho (A, B and C), RhoD, RhoE, and RhoL. These proteins and other Ras-like proteins constitute Ras superfamily of structurally and functionally related GTPase proteins. These proteins are involved in diverse physiological functions such as control of cell shape (reviewed in Tapon and Hall, *Curr. Opin. Cell Biol.* 9: 86-92 [1997]), cell motility (Aepfelbacher *et al., Proc. Natl. Acad. Sci. USA* 91: 4263-4267 [1994]; and *Curr. Biol.* 6: 70-75 [1996]), cell polarity (Adams *et al., J. Cell Biol.* 111: 131-142 [1990]), smooth muscle contraction (Hirata *et al., J. Biol. Chem.* 267: 8719-8722 [1992]), cell

adhesion (Nobes and Hall, *Cell* 81: 53-62 [1995]; Braga *et al., J. Cell Biol.* 137: 1421-1431 [1997]), cell division (Dutartre *et al., J. Cell Sci.* 109:367-377 [1996]), vesicular transport between organelles such as receptor-mediated endocytosis (Lamaze *et al., Nature* 382: 177-179 [1996]), apoptosis (Esteve *et al.,Oncogene* 11: 2657-2665 [1995]; Jimenez *et al., Oncogene* 10: 811-816 [1995]; Gulbins *et al., J. Biol. Chem.* 271: 26389-26394 [1996]; Moorman *et al., J. Immunol.* 156: 4146-4153 [1996]; Brenner *et al., J. Biol. Chem.* 272: 22173-22181 [1997]) and normal and pathological cell proliferation (Olson *et al., Science* 269: 1270-1272 [1995]; Hirai *et al., J. Biol. Chem.* 272: 13-16 [1997]; Khosravi-Far *et al., Mol. Cell. Biol.* 16: 3923-3933 [1996]; Qiu *et al., Mol. Cell. Biol.* 17: 3449-3458 [1997]; Roux *et al., Curr. Biol.* 7: 629-637 [1997]).

The presence of a GEF domain and protein kinase domains along with its homology to Trio suggests that H19G5 may possess guanine nucleotide exchange factor activity and protein kinase activity, both of which are shared by a number of proteins involved in signal transduction.

Expression pattern of 19G5 in rat and human tissues

Northern blot analysis revealed that R19G5 gene is highly expressed in heart as a 2 kb and a 4.4 kb transcripts (Figure 8). Hybridization was performed using multiple tissue Northern blot (Clontech, Palo Alto, CA) and ExpressHyb solution following the manufacturer's protocol. The R19G5 probe also hybridized to mRNAs from skeletal muscle and detected multiple weak bands. High background observed in a lane corresponding to skeletal muscles is not due to RNA degradation as probing of the same blot with β-actin probe detected the right sized transcripts with a clean background (Figure 8). The significance of the high background in skeletal muscle is not clear. Expression of R19G5 was not detected in brain, kidney, spleen, lung, liver, and testis.

Hybridization of human multiple tissue Northern blot (Clontech, Palo Alto, CA) with H19G5 probe detected a strong transcript of about 5 kb and three minor transcripts of about 3 kb, 2.4 kb, and 1.8 kb in heart tissue (Figure 9). There were multiple transcripts and high background in skeletal muscle as seen with rat tissue. Only the 3 kb transcript was detected in human fetal heart. There is a low level expression of the 2.4 kb transcript in brain. H19G5 expression was not detected in spleen, lung, liver, kidney, pancreas, thymus, prostates, testis, ovary, small intestine, colon, peripheral blood leukocyte, stomach, thyroid, spinal cord, lymph node, trachea, adrenal gland, bone marrow, uterus muscle, or bladder muscle by Northern blot.

In order to determine whether the multiple transcripts of 19G5 in human heart are derived from alternative splicing or multiple gene copies, Southern blot analysis of genomic DNA was performed with H19G5 probe (Figure 10). The results suggest that H19G5 is a single copy gene. This is consistent with the lack of detection of polymorphism in the restriction analysis of the genomic clones of H19G5.

Functional characterization of H19G5 protein

A 20 amino acid peptide from the C-terminus of H19G5 protein (SEQ ID NO: 1) was used to raise polyclonal antibodies in rabbits. Affinity purified rabbit anti-H19G5 antiserum has been obtained. Two different GST-H19G5 fusion proteins were used as antigens to generate mouse monoclonal anti-H19G5 antibodies. One of the antigens used was a GST-H19G5 fusion protein containing amino acid residues 610 to 811 of SEQ ID NO: 1. Many clones of anti-19G5 monoclonal antibodies were obtained. Three of them were shown to recognize recombinant H19G5 proteins expressed in mammalian cells by Western analysis and immunocytochemistry. The antibodies may be used to determine the size and localization of 19G5 protein by Western blot and immunohistochemistry.

Subcellular localization of H19G5 protein was determined using confocal microscopy on cells transfected with vectors expressing 19G5 proteins fused to Green Fluorescent Protein (GFP). Three 19G5-GFP fusion protein expression constructs were made using three different 19G5 cDNA clones, the longest human 19G5 clone C11 [H19G5(C11)-GFP], a 2.7 kb clone of human 19G5 containing the C-terminal kinase domain [H19G59F1)-GFP], and the rat 19G5 small transcript [R19G5(S)-GFP]. The control GFP vector and the 19G5-GFP fusion expression constructs were transfected into C2C12 cells (ATCC Catalog No. CRL-1772). Cells were grown on chamber slides and transfected using SuperFectamine reagent from Quiagen for 3 hrs. Cells were fixed 24 hrs post-transfection with 4% paraformaldehyde for 15 min at room temperature and examined using a confocal microscope. The two longer forms of human 19G5-GFP proteins were detected in the nuclei whereas the short rat 19G5-GFP fusion protein was detected in the cytoplasm (Figure 11). This suggests that different forms of 19G5 proteins are localized in different regions in the cells and may have different functions. Since the gene is normally expressed in heart, the localization of 19G5 protein in cardiac myocytes may be examined. Antisense constructs to study the effects of inhibiting the activity of 19G5 protein on cardiac myocytes may also be made and used.

A possibility that H19G5 might interact with various G proteins, as suggested by the presence of GEF domain and homology with Trio, was examined. HEK 293 cells constitutively expressing Epstein-Barr Virus Nuclear Antigen (EBNA) (Invitrogen, San Diego, CA) were transfected with H19G5-C11 expression construct. The eukaryotic vector used for 19G5 expression, pEAK8 (EdgeBiosystems), contains Epstein-Barr virus (EBV) origin of DNA replication, which allows replication of the expression construct in transfected cells thereby amplifying the level of protein expression. After 24 hrs, one 10 cm dish of transfected 293 EBNA cells were lysed with 1 ml of lysis buffer (1X PBS, 0.1% Triton, and proteinase inhibitors (0.2 mM AEBSF, 0.16 μM Aprotinin, 0.01 mM Bestatin, 3 μM E-64, 4 μM Leupeptin, and 2 μM Pepstatin) on ice for 30'. Cells were then homogenized with a dounce homogenizer on ice. The lysates were cleared by centrifugation. One ml of the cleared lysate was incubated with 5 μg of various small G proteins, expressed as GST fusion proteins and bound to glutathione-agarose beads, at 4° C for 2 hrs. The protein-agarose complex was pelleted by brief centrifugation and washed for 4 times 5' each with the lysis buffer at room temperature. The complex was then resuspended in SDS sample buffer, boiled and run on a SDS-PAGE and Western blotted with an anti-H19G5 monoclonal antibody. As shown in Figure 12, H19G5 was found to bind to Cdc42. No binding was detected with Rac1 or RhoA. Lane 8 is His-tagged Cdc42 and in lane 9 His-tagged Cdc42-agarose

complex was boiled for 5' before adding to the H19G5 cell lysate. This result showed that H19G5 specifically interacts with Cdc42 protein and there is no non-specific interaction with agarose beads. The results presented herein suggest that 19G5 protein may play an important role, by virtue of regulating a small GTPase such as Cdc42Hs, in a variety of cellular activities. For example, Cdc42 has been shown to regulate actin polymerization and focal adhesion complex formation which in turn is necessary for filopedia formation (Nobes and Hall, *Cell* 81: 53-62 [1995]). Cdc42 and rac have also been shown to regulate Jun N-terminal kinase (JNK) activity via the MAP kinase pathway (Coso *et al., Cell* 81: 1137-1146 [1995]; Minden *et al., Cell* 81: 1147-1157 [1995]; Olson *et al., Science* 269: 1270-1272 [1995]), an evolutionarily conserved and ubiquitous signal transduction pathway that impacts upon a number of important cellular functions.

C2C12 myoblasts cells (ATCC Catalog No. CRL-1772) can be induced to differentiate into myotubes when placed in a medium containing 2% horse serum (Lechner *et al., Proc. Natl. Acad. Sci. USA* 93: 4355-4359 [1996]). Figure 13 shows phase contrast micrographs of C2C12 myoblasts undergoing differentiation into myotubes. At Day 0, undifferentiated C2C12 cells with typical myoblast morphology can be seen. Once induced to differentiate, as shown here at Day 3 and 5 after induction, an increasingly larger number of cells with typical morphology of differentiated myotubes, i.e. large, elongated, multinucleated syncytial cells, could be seen. The induction of Myogenin expression was monitored during differentiation. Myogenin is not expressed in myoblasts, however, its expression is strongly induced when myoblasts undergo differentiation into myotubes. Thus, it acts as a biochemical marker of myotubes. Figure 14 shows induction of myogenin protein expression when C2C12 myoblast cells are induced to differentiate into myotubes. C2C12 cells were cultured and induced to differentiate by placing in a medium containing 2% horse serum. Cells were lysed in radioimmunoprecipitation (RIPA) buffer (1X PBS containing 1% Igepal CA-630, 0.5% sodium deoxycholate and 0.1% SDS) on ice for 30'. Total lysates were cleared by centrifugation at 10,000 rpm for 10'. Protein concentration of each lysate was measure using the BCA method. SDS sample buffer was added to the total lysate and boiled for 3'. Equal amount of total protein of each sample was run on SDS-PAGE and blotted using an anti-myogenin monoclonal antibody. The result demonstrated that Myogenin expression was induced concomitantly with differentiation of C2C12 myoblasts into myotubes under the conditions used for induction.

The expression of 19G5 RNA was also monitored at various stages during differentiation of myoblasts into myotubes. Undifferentiated C2C12 cells were plated on 6 cm dishes at 2.4 x 10^4 cells/cm^2 and cultured in growth medium with 10% fetal bovine serum for 24 hrs to about confluence. Cells were washed with PBS and induced to differentiate into myotubes in differentiation medium with 2% horse serum. Total RNA was isolated from cells at 1, 2, or 4 days post-induction using Qiagen's Rneasy kit according to the manufacture's instruction. The expression of 19G5 transcript was analyzed using Taqman assay. One microgram of total RNA, isolated at various time points post-induction, was reverse transcribed into cDNA using PE Biosystems Reagents and Multiscribe enzyme according to manufacture's instruction. Ten ng of cDNA was added to 1X master mix, and the primers and probe for the gene of interest were added according to manufacture's instructions. The reaction were carried out in the ABI Prism 7700

Detection System. The quantity of 19G5 and the quantity for 18S were determined for each sample, and the ratio of 19G5/ 18S was used to evaluate differences in the level of 19G5 expression in various samples. Comparison of the values thus obtained with pre-induction values allowed to determine fold induction of 19G5 expression during differentiation. As shown in Figure 15, expression of 19G5 transcript was significantly increased during differentiation of C2C12 myoblasts into myotubes. The level of induction reached to about 10-fold at 4 days post-induction. Increased expression of 19G5 likely reflects a specific function in myotubes.

TGF-β is known to inhibit differentiation of C2C12 myoblast into myotubes (Katagiri *et al., J. Cell Biol.* 127: 1755-1766 [1994]; Namiki *et al., J. Biol. Chem.* 272: 22046-22052 [1997]). Therefore, the effect of TGF-β on the induction of expression of 19G5 during differentiation of C2C12 cells was examined. C2C12 cells were plated in 6 cm dishes at 2.4 x 10^4 cells/cm^2 and cultured in the growth medium for 24 hours. Cells were then rinsed with PBS and induced to differentiate for 4 days in the medium containing 5% fetal bovine serum either in the absence or in the presence of 10 ng/ml of TGF-β. Total RNA was isolated from cells using Qiagen's Rneasy kit. Induction of 19G5 expression in TGF-β treated or untreated cells over undifferentiated C2C12 cells was measured by Taqman assay. As shown in Figure 16, the induction of 19G5 expression during C2C12 differentiation is inhibited by TFG-β. TGF-β is a powerful regulator of cell growth and differentiation and regulation of expression of 19G5 by TGF-β likely represents an important physiological event with significant relevance to normal and abnormal changes in cardiac cells. Furthermore, the results suggest a possible inolvement of 19G5 in myogenesis.

Example 2 (Polypeptides which can mediate signal transduction)

The polypeptides of the present invention, such as the specific embodiment shown in SEQ ID NOs: 1, 4, 6, 7, 8 or 9 may be prepared by any known techniques. Conveniently, the polypeptides may be prepared using the solid-phase synthetic technique initially described by Merrifield in *J. Am. Chem. Soc.* 15:2149-2154 (1963). Other peptide synthesis techniques may be found, for example, in M. Bodanszky *et al.*, PEPTIDE SYNTHESIS, John Wiley & Sons, 2d Ed. (1976) as well as in other reference works known to those skilled in the art. A summary of peptide synthesis techniques may be found in J. Stuart and J.D. Young, SOLID PHASE PEPTIDE SYNTHELIA, Pierce Chemical Co., Rockford, IL (1984). The synthesis of peptides by solution methods may also be used, as described in THE PROTEINS, Vol-II, 3d Ed., Neurath, H. *et al.*, Eds., p.105-237, Academic Press, New York, NY (1976). Appropriate protective groups for use in such syntheses will be found in the above texts as well as in J.F.W. McOmie, PROTECTIVE GROUPS IN ORGANIC CHEMISTRY, Plenum Press, New York, NY (1973). In general, these synthetic methods involve the sequential addition of one or more amino acid residues or suitably protected amino acid residues to a growing peptide chain. Normally, either the amino or carboxyl group of the first amino acid residue is protected by a suitable, selectively-removable protecting group. A different, selectively-removable protecting group is utilized for amino acids containing a reactive side group, such as lysine.

Using a solid phase synthesis as an example, the protected or derivatized amino acid is attached to an inert solid support through its unprotected carboxyl or amino group. The protecting group of the amino or carboxyl group is then selectively removed and the next amino acid in the sequence having the complementary (amino or carboxyl) group suitably protected is admixed and reacted under conditions suitable for forming the amide linkage with the residue already attached to the solid support. The protecting group of the amino or carboxyl group is then removed from this newly added amino acid residue, and the next amino acid (suitably protected) is then added, and so forth. After all the desired amino acids have been linked in the proper sequence, any remaining terminal and side group protecting groups (and solid support) are removed sequentially or concurrently, to provide the final peptide. The polypeptides of the invention preferably are devoid of benzylated or methylbenzylated amino acids. Such protecting group moieties may be used in the course of synthesis, but they are removed before the polypeptides are used. Additional reactions may be necessary, as described elsewhere to form intramolecular linkages to restrain conformation, if desired. The polypeptides of the present invention may also be linked to an additional sequence of amino acids either or both at the N-terminus and at the C-terminus. Such additional amino acid sequences, or linker sequences, can be conveniently affixed to a detectable label, solid matrix, or carrier. Typical amino acid residues used for linking are tyrosine, cysteine, lysine, glutamic acid and aspartic acid, or the like.

Of course, the present polypeptides may also be prepared by recombinant DNA techniques as described, for example, in Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL, Chapters 1-18, Second Edition (Cold Spring Harbor NY 1989), and as detailed in Examples 3-4 *infra*. The present invention also relates to vectors comprising DNA molecules of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques. Host cells may be genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are preferably those previously used with the host cell selected for expression, and will be apparent to the skilled artisan.

Example 3 (Hosts, Vectors and DNA encoding polypeptides which can mediate signal transduction)

The DNA molecules of the present invention may be employed for producing the polypeptides of the present invention by recombinant techniques. More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences broadly described above. Thus, for example, the DNA molecule sequence may be included in any one of a variety of expression vehicles, in particular vectors or plasmids for expressing such a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, *e.g.*, derivatives of SV40; bacterial plasmids; phage DNA; yeast plasmids; vectors derived from combinations of plasmids and phage DNA;

viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, phagescript, psiX174, pBluescript SK, pBsKS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTRC99A, pKK223-3, pKK233-3, pDR540, PRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, pXT1, pSG (Stratagene); pSVK3, pBPV, pMSG, PSVL (Pharmacia). However, any other vector or plasmid may be used as long as they are replicable and viable in the host. The vector containing the appropriate DNA sequence, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the polypeptides of the present invention. Representative examples of appropriate hosts include: bacterial cells, such as *E. coli, Salmonella typhimurium, Streptomyces*; fungal cells, such as yeast; insect cells, such as Drosophila S2 and Spodoptera Sf9; animal cells, such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector may be operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. Suitable promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, $P_L$and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of skill in the art. The expression vector may also contain a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably may contain a gene to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

An embodiment of the invention is an isolated DNA molecule comprising the nucleotide sequence of SEQ ID NOs: 2, 3 or 5. This nucleotide sequence, or fragments or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of the polypeptides of the present invention, or functionally active peptides or functional equivalents thereof, in appropriate host cells. Due to the degeneracy of the nucleotide coding sequence, other DNA sequences which encode substantially the same amino acid sequences as depicted in SEQ ID NOS: 1, 4, 6, 7, 8 or 9, or analogs or fragments thereof, may be used in the practice of the invention for the cloning and expression of a mediator of signal transduction. Such alterations include deletions, additions or substitutions of different nucleotide residues resulting in a sequence that encodes the same or a functionally equivalent gene product. The gene product may contain deletions, additions or substitutions of amino acid residues within the sequence, which

result in a silent change thus producing a bioactive product. Such amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, the amphipathic nature of the residues involved and/or on the basis of crystallographic data. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; amino acids with uncharged polar head groups having similar hydrophilicity values include the following: leucine, isoleucine, valine; glycine, alanine; asparagine, glutamine; serine, threonine; phenylalanine, tyrosine.

Techniques well known to those skilled in the art for the isolation of DNA, generation of appropriate restriction fragments, construction of clones and libraries, and screening recombinants may be used. For a review of such techniques, see, for example, Sambrook, *et al., supra*, the disclosure of which is hereby incorporated by reference. Also, the 5' untranslated and coding regions of the nucleotide sequence could be altered to improve the translational efficiency of the mRNA. In addition, based on X-ray crystallographic data, sequence alterations could be undertaken to improve protein stability, *e.g.*, introducing disulfide bridges at the appropriate positions, and/or deleting or replacing amino acids that are predicted to cause protein instability. These are only examples of modifications that can be engineered to produce a more active or stable protein, more protein, or even change the substrate specificity of the protein.

Example 4 (Cells transformed with recombinant vectors containing DNA encoding polypeptides which can mediate signal transduction)

In a further embodiment, the present invention relates to host cells containing the above-described construct. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. The host cell preferably may secrete the recombinant protein. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (L. Davis *et al.*, BASIC METHODS IN MOLECULAR BIOLOGY, 1986)).

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, *et al., supra*.

Transcription of a DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually from about 10 to 300 bp, that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the

late side of the replication origin (base pair 100 to 270), a cytomegalovirus early promoter enhancer, a polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), alpha factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is preferably assembled in appropriate phase with translation, initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use may be constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation, initiation and termination signals in operable reading phase with a functional promoter. The vector may comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI). These pBR322 backbone sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter may be de-repressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells may be cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents. Various mammalian cell culture systems can also be employed to express recombinant polypeptides. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell*, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors may comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences,

and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The polypeptides of the present invention may be recovered and purified from recombinant cell cultures by methods used heretofore, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxyapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated with mammalian or other eukaryotic carbohydrates or may be non-glycosylated. Polypeptides of the invention may also include an initial methionine amino acid residue.

In particular, two baculovirus expression constructs of a wild-type and a mutant H19G5 C-terminal kinase domain have been constructed. The wild-type kinase domain construct was made by cloning a cDNA fragment which encodes the amino acid residues 1002 to 1314 of the H19G5 contig protein sequence into the pFastBac HTc vector. The mutant kinase domain construct contains the same amino acid sequence as the wild-type kinase domain except that Tyr residue at position 1213 was changed to a Glu in an attempt to create a constitutively active kinase. Both recombinant proteins contain a his-tag at the N-terminus.

Example 5 (Pharmaceutically acceptable salts of polypeptides which can mediate signal transduction)

Any peptide of the present invention may be used in the form of a pharmaceutically acceptable salt. Suitable acids which are capable of forming salts with the peptides of the present invention include inorganic acids such as hydrochloric acid, hydrobromic acid, perchloric acid, nitric acid, thiocyanic acid, sulfuric acid, phosphoric acid and the like; and organic acids such as formic acid, acetic acid, propionic acid, glycolic acid, lactic acid, pyruvic acid, oxalic acid, malonic acid, succinic acid, maleic acid, fumaric acid, anthranilic acid, cinnamic acid, naphthalene sulfonic acid, sulfanilic acid, and the like.

Suitable bases capable of forming salts with the peptides of the present invention include inorganic bases such as sodium hydroxide, ammonium hydroxide, potassium hydroxide and the like; and organic bases such as mono-, di- and tri-alkyl and aryl amines (*e.g.*, triethylamine, diisopropyl amine, methyl amine, dimethyl amine and the like) and optionally substituted ethanolamines (*e.g.*, ethanolamine, diethanolamine and the like).

Example 6 (Pharmaceutical compositions containing polypeptides which can mediate signal transduction)

For use in a method of identification, prevention or treatment, such as the identification, prevention or treatment of infection of a mammalian host by a microorganism, the polypeptides of the present invention may be present in a pharmaceutical composition in admixture with a pharmaceutically acceptable sterile vehicle. The pharmaceutical composition may be compounded according to conventional pharmaceutical formulation techniques.

The vehicle may take a wide variety of forms depending on the form of preparation desired for administration, *e.g.*, sublingual, rectal, nasal, oral or parenteral. Compositions for oral dosage form may include any of the usual pharmaceutical media, such as, for example, water, oils, alcohols, flavoring agents, preservatives, coloring agents and the like in the case of oral liquid preparations (*e.g.*, suspensions, elixirs and solutions) or carriers such as starches, sugars, diluents, granulating agents, lubricants, binders, disintegrating agents and the like in the case of oral solid preparations (*e.g.*, powders, capsules and tablets). Controlled release forms may also be used. Because of their ease in administration, tablets and capsules represent an advantageous oral dosage unit form, in which case solid pharmaceutical carriers may be employed. If desired, tablets may be sugar coated or enteric coated by standard techniques.

For compositions to be administered parenterally, the carrier will usually comprise sterile water, although other ingredients to aid solubility or for preservation purposes may be included. Injectable suspensions may also be prepared, in which case appropriate liquid carriers, suspending agents and the like may be employed. The parenteral routes of administration may be intravenous injection, intramuscular injection or subcutaneous injection.

For intravenous administration, the polypeptides may be dissolved in an appropriate intravenous delivery vehicle containing physiologically compatible substances such as sodium chloride, glycine and the like, having a buffered pH compatible with physiologic conditions. Such intravenous delivery vehicles are known to those skilled in the art.

The polypeptides of the invention may be administered to subjects where mediation of signal transduction is desired. The peptides may be administered by any convenient means that will result in the delivery to the subject of an effective amount to mediate signal transduction. Oral administration is presently contemplated as a preferred administration route. The amount administered will depend on the activity of the particular compound administered, which may readily be determined by those of ordinary skill in the art.

Example 7 (Monoclonal antibodies against polypeptides which can mediate signal transduction)

Another embodiment of the present invention relates to a monoclonal antibody to the polypeptides of the present invention (or an antigenic portion thereof), which may be produced by methods recognized in the art, including the formation of monoclonal antibody-producing hybridomas (Kohler, G., and C. Milstein, *Nature* 256:495-497 (1975); *Eur. J. Immunol.* 6:511-519 (1976)). By fusing antibody-forming cells (spleen lymphocytes) with myeloma cells (malignant cells of bone marrow primary tumors), a hybrid cell line is created from a single fused cell hybrid (called a

hybridoma or clone) having certain inherited characteristics of both the lymphocytes and myeloma cell lines. Like the lymphocytes (taken from animals primed with sheep red blood cells as antigen), the hybridomas secreted a single type of immunoglobulin specific to the antigen; moreover, like the myeloma cells, the hybrid cells had the potential for indefinite cell division. The combination of these two features offered distinct advantages over conventional antisera. Whereas antisera derived from vaccinated animals are variable mixtures of polyclonal antibodies which never can be reproduced identically, monoclonal antibodies are highly specific immunoglobulins of a single type. The single type of immunoglobulin secreted by a hybridoma is specific to one and only one antigenic determinant, or epitope, on the antigen, a complex molecule having a multiplicity of antigenic determinants. For instance, if the antigen is a protein, an antigenic determinant may be one of the many peptide sequences (generally 6-7 amino acids in length (Atassi, M.Z., *Molec. Cell. Biochem.* 32:21-43 (1980)) within the entire protein molecule. Hence, monoclonal antibodies raised against a single antigen may be distinct from each other depending on the determinant that induced their formation; but for any given clone, all of the antibodies it produces are identical. Furthermore, the hybridoma cell line can be reproduced indefinitely, is easily propagated *in vitro* or *in vivo*, and yields monoclonal antibodies in extremely high concentration.

Example 8 (Therapeutic monoclonal antibodies against polypeptides which can mediate signal transduction)

The monoclonal antibodies of the present invention can have potential immunotherapeutic value (Oldham, R.K., *J. Clin. Oncol.*, 1:582-590 (1983); Miller, R.A. *et al., Blood*, 62:988-995 (1983); Miller R.A. *et al., New Engl. J. Med.* 306:517-522 (1982); Ritz, J. and Schlossman, S., *Blood*, 59:1-11 (1982); and Kirch, M.E. and Ulrich, H., *J. Immunol.* 127:805-810 (1981) (investigating the therapeutic efficacy in both animal and human subjects)). In addition, the monoclonal antibodies can be used in cytotoxic drug-antibody conjugates similar to those described in Beverly, P.C.L., *Nature*, 297:358-9 (1982); Krolick, K.A. *et al., Nature*, 295:604-5 (1982); Krolick, K.A. *et al., Proc. Natl. Acad. Sci. U.S.A.*, 77:5419-23 (1980); Arnon, R. and Sela, M., *Immunol. Rev.*, 62:5-27 (1982); Raso, V. *et al., Cancer Res.*, 42:457-64 (1982); and DeWeger, R.A. and Dullens, H.F.J., *Immunol. Rev.* 62:29-45 (1982).

In an embodiment of the invention, purified polypeptides of the present invention (or an antigenic portion thereof) can be used as an antigen or immunogen. In addition, microorganisms expressing H1965 protein or polypeptide fragments thereof also represent potential antigens or sources of antigen with which to immunize animals to obtain somatic cells for fusion. Somatic cells with the potential for producing antibody and, in particular, B lymphocytes, are suitable for fusion with a B-cell myeloma line. Those antibody-producing cells that are in the dividing plasmablast stage fuse preferentially. Somatic cells may be derived from the lymph nodes, spleens and peripheral blood of primed animals and the lymphatic cells of choice depending to a large extent on their empirical usefulness in the particular fusion system. Once-primed or hyperimmunized animals can be used as a source of antibody-producing lymphocytes. Mouse lymphocytes give a higher percentage of stable fusions with mouse myeloma lines. However, the use of rat, rabbit, and frog cells is also possible. Alternatively, human somatic cells capable of producing antibody,

specifically B lymphocytes, are suitable for fusion with myeloma cell lines. While B lymphocytes from biopsied spleens or lymph nodes of individual may be used, the more easily accessible peripheral blood B lymphocytes are preferred. The lymphocytes may be derived from patients with diagnosed carcinomas.

Specialized myeloma cell lines have been developed from lymphocyte tumors for use in hybridoma-producing fusion procedures (Kohler, G., and C. Milstein, *Eur. J. Immunol.* 6:511-519 (1976); M. Schulman *et al.*, *Nature* 276: 269-270 (1978)). Examples of myeloma cell lines that may be used for the production of fused cell hybrids include X63-Ag8, NSI-Ag4/1, MPCII-45.6TGI.7, C63-Ag8.653, Sp2/0-Ag14, FO, and S194/5XX0.BU.1, all derived from mice; 210.RCY3.Agl.2.3, U-226AR, and GM1500GTGAL2, all derived from rats; and U-226AR and GM1500GTGAL2, derived from humans, (G.J. Hammerling, U. Hammerling, and J.F. Kearney (editors), *Monoclonal Antibodies and T-cell Hybridomas* in: J.L. Turk (editor) RESEARCH MONOGRAPHS IN IMMUNOLOGY, Vol. 3, Elsevier/North Holland Biomedical Press, NY (1981)).

Methods for generating hybrids of antibody-producing spleen or lymph node cells and myeloma cells usually comprise mixing somatic cells with myeloma cells in a 2:1 proportion (though the proportion may vary from about 20:1 to about 1:1), respectively, in the presence of an agent or agents (chemical or electrical) that promote the fusion of cell membranes. It is often preferred that the same species of animal serve as the source of the somatic and myeloma cells used in the fusion procedure. Fusion methods have been described by Kohler and Milstein (*Nature* 256:495-497 (1975) and *Eur. J. Immunol.* 6:511-519 (1976), and by Gefter *et al.* (*Somatic Cell Genet.* 3:231-236 (1977)). The fusion-promotion agents used by those investigators were Sendai virus and polyethylene glycol (PEG), respectively.

Once the desired fused cell hybrids have been selected and cloned into individual antibody-producing cell lines, each cell line may be propagated in either of two standard ways. A sample of the hybridoma can be injected into a histocompatible animal of the type that was used to provide the somatic and myeloma cells for the original fusion. The injected animal develops tumors secreting the specific monoclonal antibody produced by the fused cell hybrid. The body fluids of the animal, such as serum or ascites fluid, can be tapped to provide monoclonal antibodies in high concentration. Alternatively, the individual cell lines may be propagated *in vitro* in laboratory culture vessels; the culture medium, also containing high concentrations of a single specific monoclonal antibody, can be harvested by decantation, filtration or centrifugation.

Example 9 (Diagnostic monoclonal antibodies against polypeptides which can mediate signal transduction)

The monoclonal antibodies of this invention can be used as probes in detecting discrete antigens expressed by tissue or cell samples. The expression or lack of expression of these antigens can provide clinically exploitable information that is not apparent after standard histopathological evaluations. It may thus be possible to correlate the immuno-phenotypes of individual tissue or cell samples with various aspects of signal transduction and responsiveness to certain types of therapies, thereby establishing important classifications of prognosis.

The use of the monoclonal antibodies described herein can be extended to the screening of human biological fluids for the presence of the specific antigenic determinant recognized. *In vitro* immunoserological evaluation of sera withdrawn from patients thereby permits non-invasive diagnosis of disease. By way of illustration, human fluids, such as pleural fluids or lymph, can be taken from a patient and assayed for the specific epitope, either as released antigen or membrane-bound on cells in the sample fluid, using monoclonal antibodies against the polypeptides of the present invention in standard radioimmunoassays or enzyme-linked immunoassays known in the art or competitive binding enzyme-linked immunoassays.

The monoclonal antibodies of this invention are potentially useful for targeting diseased tissue or cells *in vivo*. They can therefore be used in humans for localization and monitoring of the microbial infection. For this application, it is preferable to use purified monoclonal antibodies. Purification of monoclonal antibodies for human administration by ammonium sulfate or sodium sulfate precipitation followed by dialysis against saline and filtration sterilization has been described by Miller *et al.* (in: HYBRIDOMAS IN CANCER DIAGNOSIS AND THERAPY, (1982), p. 134).

Alternatively, immunoaffinity chromatography techniques may be used to purify the monoclonal antibodies. The purified monoclonal antibodies can be labeled with radioactive compounds, for instance, radioactive iodine, and administered to a patient intravenously. After localization of the antibodies at the infection site, they can be detected by emission tomographical and radionuclear scanning techniques, thereby pinpointing the location of the infection. Experimental radioimmunodetection with monoclonal antibodies may occur by external scintigraphy.

Passive monoclonal serotherapy may be a potential use for the monoclonal antibodies of this invention. By way of illustration, purified anti-H19G5 monoclonal antibody is dissolved in an appropriate carrier, *e.g.*, saline, with or without human albumin, at an appropriate dosage and is administered to a patient. The monoclonal antibodies are preferably administered intravenously, *e.g.*, by continuous intravenous infusion over several hours, as in Miller *et al*, *supra*. Infusions can be administered over a period of weeks during which the anti-microbial effects are monitored.

Example 10 (Anti-idiotypic antibodies to antibodies against polypeptides which can mediate signal transduction)

In an alternate embodiment, the antibodies described herein are used to stimulate the production of corresponding anti-idiotypic antibodies. In brief, anti-idiotypic antibodies, or antiidiotypes are antibodies directed against the antigen combining region or variable region (idiotype) of another antibody. Based on Jerne's network model of idiotypic relationships (Jerne, *Ann. Immunol.* 125:373 (1974); Jerne *et al.*, *EMBO* 1:234 (1982)), immunization with an antibody molecule expressing a paratope (antigen-combining site) for a given antigen should produce a group of anti-antibodies, some of which share with the antigen a complementary structure to the paratope. Immunization with a subpopulation of antiidiotypic antibodies should in turn produce a subpopulation of antiidiotypic antibodies which bind the initial antigen. Thus, the administration of the monoclonal antibodies of the present

invention may result in a modification of the host's immune response, as the consequence of the formation of anti-idiotypic antibodies which may develop during therapy with the monoclonals.

Example 11 (Monoclonal antibody-drug conjugates)

The monoclonal antibodies of this invention can be used in conjunction with a broad spectrum of pharmaceutical or cytotoxic agents that selectively affect diseased tissue or cells over normal tissues or cells in the mammalian host. The methods used for binding the cytotoxic agents to the monoclonal antibody molecule can involve either non-covalent or covalent linkages. Since non-covalent bonds are more likely to be broken before the antibody complex reaches the target site, covalent linkages are preferred. For instance, carbodiimide can be used to link carboxy groups of the pharmaceutical agent to amino groups of the antibody molecule. Bifunctional agents such as dialdehydes or imidoesters can be used to link the amino group of a drug to amino groups of the antibody molecule. The Schiff base reaction can be used to link drugs to antibody molecules. This method involves the periodate oxidation of a drug or cytotoxic agent that contains a glycol or hydroxy group, thus forming an aldehyde that is then reacted with the antibody molecule. Attachment occurs via formation of a Schiff base with amino groups of the antibody molecule. Additionally, drugs with reactive sulfhydryl groups have been coupled to antibody molecules.

Example 12 (Diagnostic kit)

Another embodiment of the invention relates to a diagnostic kit for detecting diseased tissue or cells using an antibody against a polypeptide which can mediate signal transduction. The diagnostic kit may further comprise, where necessary, other components of the signal producing system, including agents for reducing background interference, control reagents, or an apparatus, container or other solid support for conducting the test. The binding of antibody to the target can be detected by well known methods, including radiation (*e.g.*, use of a radioactive nucleotide), colorimetry (*e.g.*, use of an enzyme that can cause a color change in a substrate), fluorescence (*e.g.*, use of a dye such as propidium iodide, fluorescein, or rhodamine), and luminescence (*e.g.*, use of an alkaline phosphatase substrate that releases photons upon cleavage or luciferin). Detection can be qualitative or quantitative.

Example 13 (Gene therapy)

Another embodiment of the present invention involves the use of the DNA of the present invention in gene therapy applications. Gene therapy has been broadly defined as "the correction of a disease phenotype through the introduction of new genetic information into the affected organism" (Roemer, K. and Friedmann, T., *Eur. J. Biochem.* 208: 211-225 (1992)). Two basic approaches to gene therapy have evolved: (1) *ex vivo* gene therapy and (2) *in vivo* gene therapy. In *ex vivo* gene therapy, cells are removed from a subject and cultured *in vitro*. A functional replacement gene is introduced into the cells (transfection) *in vitro*, the modified cells are expanded in culture, and then reimplanted in the subject. These genetically modified, reimplanted cells are reported to secrete detectable levels

of the transfected gene product *in situ* (Miller, A.D., *Blood* 76: 271-278 (1990)) and Selden, R.F., *et al., New Eng. J. Med.* 317: 1067-1076 (1987)). The development of improved retroviral gene transfer methods (transduction) facilitates the transfer into and subsequent expression of genetic material by somatic cells (Cepko, C.L., *et al., Cell* 37: 1053-1062 (1984)). Accordingly, retrovirus-mediated gene transfer has been used in clinical trials to mark autologous cells and as a way of treating genetic disease (Rosenberg, S.A., *et al., New Eng. J. Med.* 323: 570-578 (1990); Anderson, W.F., *Human Gene Therapy* 2: 99-100 (1991)). Several *ex vivo* gene therapy studies in humans are reported (reviewed in Anderson, W.F., *Science* 256: 808-813 (1992) and Miller A.D., *Nature* 357: 455-460 (1992)).

In *in vivo* gene therapy, target cells are not removed from the subject. Rather, the transferred gene is introduced into cells of the recipient organism *in situ*, that is, within the recipient. *In vivo* gene therapy has been examined in several animal models (reviewed in Felgner, P.L. and Rhodes, G., *Nature* 349: 351-352 (1991)). Publications have reported the feasibility of direct gene transfer *in situ* into organs and tissues such as muscle (Ferry, N., *et al., Proc. Natl. Acad. Sci.* 88: 8377-8781 (1991); Quantin, G., *et al., Proc. Natl. Acad. Sci. USA* 89: 2581-2584 (1992)), hematopoietic stem cells (Clapp, D.W., *et al., Blood* 78: 1132-1139 (1991)), the arterial wall (Nabel, E.G., *et al., Science* 244: 1342-1344 (1989)), the nervous system (Price, J.D., *et al., Proc. Natl. Acad. Sci.* 84: 156-160 (1987)), and lung (Rosenfeld, M.A., *et al., Science* 252: 431-434 (1991)). Direct injection of DNA into skeletal muscle (Wolff, J.A., *et al., Science* 247: 1465-1468 (1990)), heart muscle (Kitsis, R.N., *et al., Proc. Natl. Acad. Sci. USA* 88: 4138-4142 (1991)) and injection of DNA-lipid complexes into the vasculature (Lim, C.S., *et al., Circulation* 83: 2007-2011 (1991); Ledere, G.D., *et al., J. Clin. Invest.* 90: 936-944 (1992); Chapman, G.D., *et al., Circ. Res.* 71: 27-33 (1992)) also have been reported to yield a detectable expression level of the inserted gene product(s) *in vivo*.

Recent gene therapy efforts have been aimed at the identification of various cell types for transformation, including keratinocytes (Morgan, J.R., *et al., Science* 237: 1476-1479 (1987)), fibroblasts (Palmer, T.D., *et al., Proc. Natl. Acad. Sci.* 88: 1330-1334 (1991); Garver Jr., R.I., *et al., Science* 237: 762-764 (1987); International Patent Application PCT/US92/01890, having publication number WO 92/15676), lymphocytes (Reimann, J.K., *et al., J. Immunol. Methods* 89: 93-101 (1986)), myoblasts (Barr, E. and Leiden, J.M., *Science* 254: 1507-1509 (1991); Dai, Y. *et al.*, PNAS 89: 10892-10895 (1992); Roman, M., *et al., Somatic Cell and Molecular Genetics* 18: 247-258 (1992)), smooth muscle cells (Lynch, C.M. *et al., Proc. Natl. Acad. Sci. USA* 89: 1138-1142 (1992)), and epithelial cells (Nabel, E.G., *et al., Science* 244: 1342-1344 (1989)), International Patent Application PCT/US89/05575 (having publication number WO 90/06997), the contents of which references and patent/patent applications are incorporated herein by reference.

The delivery of an effective dose of a prophylactic or therapeutic agent *in situ* depends on the efficiency of transfection (or transduction) as well as the number of target cells. Epithelial cell-based gene therapy, in particular, involves a relatively small area available *in situ* for receiving genetically modified epithelial cells. The delivery of an effective dose of prophylactic or therapeutic agent *in situ* thus depends upon the total number of implanted epithelial cells.

In one embodiment of the invention, exogenous genetic material (*e.g.*, a cDNA encoding a polypeptide of the present invention) is introduced into a syngeneic host cell *ex vivo* or *in vivo* by genetic transfer methods, such as transfection or transduction, to provide a genetically modified host cell. Various expression vectors (*i.e.*, vehicles for facilitating delivery of exogenous genetic material into a target cell) are known to one skilled in the art.

Transfection refers to the insertion of nucleic acid into a mammalian host cell using physical or chemical methods. Several transfection techniques are known to those of ordinary skill in the art including: calcium phosphate DNA co-precipitation (METHODS IN MOLECULAR BIOLOGY, Vol. 7, *Gene Transfer and Expression Protocols*, Ed. E.J. Murray, Humana Press (1991)); DEAE-dextran; electroporation; cationic liposome-mediated transfection; and tungsten particle-facilitated microparticle bombardment (Johnston, S.A., *Nature* 346: 776-777 (1990)). Strontium phosphate DNA co-precipitation (Brash D.E. *et al. Molec. Cell. Biol.* 7: 2031-2034 (1987)) is a preferred transfection method.

In contrast, transduction refers to the process of transferring nucleic acid into a cell using a DNA or RNA virus. A RNA virus (*i.e.*, a retrovirus) for transferring a nucleic acid into a cell is referred to herein as a transducing chimeric retrovirus. Exogenous genetic material contained within the retrovirus is incorporated into the genome of the transduced host cell. A host cell that has been transduced with a chimeric DNA virus (*e.g.*, an adenovirus carrying a cDNA encoding a therapeutic agent) will not have the exogenous genetic material incorporated into its genome, but will be capable of expressing the exogenous genetic material that is retained extrachromosomally within the cell.

Typically, the exogenous genetic material includes the heterologous gene (usually in the form of a cDNA comprising the exons coding for the therapeutic protein) together with a promoter to control transcription of the new gene. The promoter characteristically has a specific nucleotide sequence necessary to initiate transcription. Optionally, the exogenous genetic material further includes additional sequences (*i.e.*, enhancers) required to obtain the desired gene transcription activity. For the purpose of this discussion an enhancer is simply any non-translated DNA sequence which works contiguous with the coding sequence (in *cis*) to change the basal transcription level dictated by the promoter. Preferably, the exogenous genetic material is introduced into the host cell genome immediately downstream from the promoter so that the promoter and coding sequence are operatively linked so as to permit transcription of the coding sequence. A preferred retroviral expression vector includes an exogenous promoter element to control transcription of the inserted exogenous gene. Such exogenous promoters include both constitutive and inducible promoters.

Naturally-occurring constitutive promoters control the expression of essential cell functions. As a result, a gene under the control of a constitutive promoter is expressed under all conditions of cell growth. Exemplary constitutive promoters include the promoters for the following genes which encode certain constitutive or housekeeping functions: hypoxanthine phosphoribosyl transferase (HPRT), dihydrofolate reductase (DHFR) (Scharfmann *et al., Proc. Natl. Acad. Sci. USA* 88: 4626-4630 (1991)), adenosine deaminase, phosphoglycerol kinase (PGK), pyruvate kinase, phosphoglycerol mutase, the beta -actin promoter (Lai *et al., Proc. Natl. Acad. Sci. USA* 86: 10006-10010 (1989)), and other constitutive promoters known to those of skill in the art. In addition, many viral

promoters function constitutively in eukaryotic cells. These include: the early and late promoters of SV40, the long terminal repeats (LTRs) of Moloney Leukemia Virus and other retroviruses, and the thymidine kinase promoter of Herpes Simplex Virus, among many others. Accordingly, any such constitutive promoters can be used to control transcription of a heterologous gene insert.

Genes that are under the control of inducible promoters are expressed only or to a greater degree, in the presence of an inducing agent, (*e.g.*, transcription under control of the metallothionein promoter is greatly increased in presence of certain metal ions). Inducible promoters include responsive elements (REs) which stimulate transcription when their inducing factors are bound. For example, there are REs for serum factors, steroid hormones, retinoic acid and cyclic AMP. Promoters containing a particular RE can be chosen in order to obtain an inducible response, and in some cases, the RE itself may be attached to a different promoter, thereby conferring inducibility to the recombinant gene. Thus, by selecting the appropriate promoter (constitutive versus inducible; strong versus weak), it is possible to control both the existence and level of expression of a therapeutic agent in the genetically modified host cell. If the gene encoding the prophylactic or therapeutic agent is under the control of an inducible promoter, delivery of the agent *in situ* is triggered by exposing the genetically modified cell *in situ* to conditions for permitting transcription of the prophylactic or therapeutic agent, *e.g.*, by intraperitoneal injection of specific inducers of the inducible promoters which control transcription of the agent. For example, *in situ* expression by genetically modified host cells of a therapeutic agent encoded by a gene under the control of the metallothionein promoter, is enhanced by contacting the genetically modified cells with a solution containing the appropriate (*i.e.*, inducing) metal ions *in situ*.

Accordingly, the amount of therapeutic agent that is delivered *in situ* is regulated by controlling such factors as: (1) the nature of the promoter used to direct transcription of the inserted gene (*i.e.*, whether the promoter is constitutive or inducible, strong or weak); (2) the number of copies of the exogenous gene that are inserted into the host cell; (3) the number of transduced/transfected host cells that are administered (*e.g.*, implanted) to the patient; (4) the size of the implant (*e.g.*, graft or encapsulated expression system); (5) the number of implants; (6) the length of time the transduced/transfected cells or implants are left in place; and (7) the production rate of the prophylactic or therapeutic agent by the genetically modified host cell. Selection and optimization of these factors for delivery of an effective dose of a particular prophylactic or therapeutic agent is deemed to be within the scope of one of skill in the art, taking into account the above-disclosed factors and the clinical profile of the patient.

In addition to at least one promoter and at least one heterologous nucleic acid encoding the prophylactic or therapeutic agent, the expression vector preferably includes a selection gene, for example, a neomycin resistance gene, for facilitating selection of host cells that have been transfected or transduced with the expression vector. Alternatively, the host cells are transfected with two or more expression vectors, at least one vector containing the gene(s) encoding the prophylactic or therapeutic agent(s), the other vector containing a selection gene. The selection of a suitable promoter, enhancer, selection gene and/or signal sequence is deemed to be within the scope of one skilled in the art.

The prophylactic or therapeutic agent can be targeted for delivery to an extracellular, intracellular or membrane location. If it is desirable for the gene product to be secreted from the host cells, the expression vector is designed to include an appropriate secretion signal sequence for secreting the therapeutic gene product from the cell to the extracellular milieu. If it is desirable for the gene product to be retained within the host cell, this secretion signal sequence is omitted. In a similar manner, the expression vector can be constructed to include retention signal sequences for anchoring the prophylactic or therapeutic agent within the host cell plasma membrane. For example, membrane proteins have hydrophobic transmembrane regions that stop translocation of the protein in the membrane and do not allow the protein to be secreted. The construction of an expression vector including signal sequences for targeting a gene product to a particular location is deemed to be within the scope of one of skill in the art.

In an embodiment, vectors for mammalian host cell gene therapy are viruses, more preferably replication-deficient viruses (described in detail below). Exemplary viral vectors are derived from: Harvey Sarcoma Virus; Rous Sarcoma Virus, MPSV, Moloney murine leukemia virus and DNA viruses (*e.g.*, adenovirus). *See* Temin, H., *Retrovirus vectors for gene transfer*, in GENE TRANSFER, Kucherlapati R, Ed., pp. 149-187 (Plenum 1986).

Replication-deficient retroviruses are capable of directing synthesis of virion proteins, but are incapable of making infectious particles. Accordingly, these genetically altered retroviral expression vectors have general utility for high-efficiency transduction of genes in cultured cells, and specific utility for use in the method of the present invention. Such retroviruses further have utility for the efficient transduction of genes into host cells *in vivo*. Retroviruses have been used extensively for transferring genetic material into cells. Standard protocols for producing replication-deficient retroviruses (including the steps of incorporation of exogenous genetic material into a plasmid, transfection of a packaging cell line with plasmid, production of recombinant retroviruses by the packaging cell line, collection of viral particles from tissue culture media, and infection of the target cells with the viral particles) are provided in Kriegler, M. GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, W.H. Freeman Co., NY (1990) and Murray, E.J., ed. METHODS IN MOLECULAR BIOLOGY, Vol. 7, Humana Press Inc., Clifton, NJ (1991).

The major advantage of using retroviruses for gene therapy is that the viruses insert the gene encoding the therapeutic agent into the host cell genome, thereby permitting the exogenous genetic material to be passed on to the progeny of the cell when it divides. In addition, gene promoter sequences in the LTR region have been reported to enhance expression of an inserted coding sequence in a variety of cell types (*see e.g.*, Hilberg *et al.*, *Proc. Natl. Acad. Sci. USA* 84: 5232-5236 (1987); Holland *et al.*, *Proc. Natl. Acad. Sci. USA* 84: 8662-8666 (1987); Valerio *et al.*, *Gene* 84: 419-427 (1989)). *In vivo* gene therapy using replication-deficient retroviral vectors to deliver a therapeutically effective amount of a therapeutic agent can be efficacious if the efficiency of transduction is high and/or the number of target cells available for transduction is high.

Yet another viral candidate useful as an expression vector for transformation of mammalian host cells is the adenovirus, a double-stranded DNA virus. The adenovirus is frequently responsible for respiratory tract infections in humans and thus appears to have an avidity for the epithelium of the respiratory tract (Straus, S., THE ADENOVIRUS,

H.S. Ginsberg, Editor, Plenum Press, NY, p.451-496 (1984)). Moreover, the adenovirus is infective in a wide range of cell types, including, for example, muscle and epithelial cells (Larrick, J.W. and Burck, K.L., GENE THERAPY. APPLICATION OF MOLECULAR BIOLOGY, Elsevier Science Publishing Co., Inc., NY, p.71-104 (1991)). The adenovirus also has been used as an expression vector in muscle cells *in vivo* (Quantin, B., *et al., Proc. Natl. Acad. Sci. USA* 89: 2581-2584 (1992)).

Like the retrovirus, the adenovirus genome is adaptable for use as an expression vector for gene therapy, *i.e.*, by removing the genetic information that controls production of the virus itself (Rosenfeld, M.A., *et al., Science* 252:431-434 (1991)). Because the adenovirus functions in an extrachromosomal fashion, the recombinant adenovirus does not have the theoretical problem of insertional mutagenesis.

Thus, as will be apparent to one skilled in the art, a variety of suitable viral expression vectors are available for transferring exogenous genetic material into mammalian host cells. The selection of an appropriate expression vector to express an agent for the identification, prevention or treatment of microbial infection amenable to gene replacement therapy and the optimization of the conditions for insertion of the selected expression vector into the cell are within the scope of one of skill in the art without the need for undue experimentation.

In an alternative embodiment, the expression vector is in the form of a plasmid, which is transferred into the target host cells by one of a variety of methods: physical (*e.g.*, microinjection (Capecchi, M.R., *Cell* 22: 479-488 (1980)), electroporation (Andreason, G.L. and Evans, G.A. *Biotechniques* 6: 650-660 (1988)), scrape loading, microparticle bombardment (Johnston, S.A., *Nature* 346: 776-777 (1990)) or by cellular uptake as a chemical complex (*e.g.*, calcium or strontium co-precipitation, complexation with lipid, complexation with ligand) (METHODS IN MOLECULAR BIOLOGY, Vol. 7, GENE TRANSFER AND EXPRESSION PROTOCOLS, Ed. E. J. Murray, Humana Press (1991)). Several commercial products are available for cationic liposome complexation including Lipofectin (Life Technologies, Inc., Gaithersburg, MD) (Felgner, P.L., *et al., Proc. Natl. Acad. Sci.* 84: 7413-7417 (1987)) and Transfectam™ (ProMega, Madison, Wis.) (Behr, J.P., *et al., Proc. Natl. Acad. Sci. USA* 86: 6982-6986 (1989); Loeffler, J.P., *et al., J. Neurochem.* 54: 1812-1815 (1990)). However, the efficiency of transfection by these methods is highly dependent on the nature of the target cell and accordingly, the conditions for optimal transfection of nucleic acids into host cells using the above-mentioned procedures must be optimized. Such optimization is within the scope of one of skill in the art.

In an embodiment, the preparation of genetically modified host cells contains an amount of cells sufficient to deliver a prophylactically or therapeutically effective dose of a signal transduction mediator of the present invention to the recipient *in situ*. The determination of an effective dose of the prophylactic or therapeutic agent for a known microbial infection is within the scope of one of skill in the art. Thus, in determining the effective dose, the skilled artisan would consider the condition of the patient, the severity of the condition, as well as the results of clinical studies of the prophylactic or therapeutic agent being administered.

If the genetically modified host cells are not already present in a pharmaceutically acceptable carrier, they are placed in such a carrier prior to administration to the recipient. Such pharmaceutically acceptable carriers include, for example, isotonic saline and other buffers as appropriate to the patient and therapy. The genetically modified cells are administered by, for example, intraperitoneal injecting or implanting the cells or a graft or capsule containing the cells in a host cell-compatible site of the recipient. As used herein, host cell-compatible site refers to a structure, cavity or fluid of the recipient into which the genetically modified cell(s), host cell graft, or encapsulated host cell expression system can be implanted, without triggering adverse physiological consequences. Representative host cell-compatible sites include, for example, the peritoneal, pleural and pericardial cavities. Preferably, the host cell-compatible site communicates with the lymphatic system, thereby enabling delivery of the therapeutic agent to the vascular system.

In one embodiment, the host cell-compatible site may be denuded prior to implanting the cells. Exemplary denuding methods include but are not limited to: (1) injection of distilled water into the site (*e.g.*, the peritoneal cavity) for 20 minutes, followed by scraping off a portion of the epithelial layer; (2) injection of 0.1% buffered trypsin for 20 minutes followed by scraping; (3) removal of epithelial cells by gentle scraping with a cell scraper and (4) touching a piece of Gelfilm (Upjohn, Kalamazoo, MI) to the endothelium.

The genetically modified host cells are implanted in a host cell-compatible site, alone or in combination with other genetically modified host cells. Thus, the instant invention embraces a method for modifying the epithelial system of a recipient by using a mixture of genetically modified host cells, such that a first modified cell expresses a first prophylactic or therapeutic agent of the present invention and a second modified cell expresses a second prophylactic or therapeutic agent. Other genetically modified cell types (*e.g.*, hepatocytes, smooth muscle cells, fibroblasts, glial cells, mesothelial cells or keratinocytes) can be added, together with the genetically altered epithelial cells, to produce expression of a complex set of introduced genes. Moreover, more than one recombinant gene can be introduced into each genetically modified cell on the same or different vectors, thereby allowing the expression of multiple prophylactic or therapeutic agents of the present invention by a single cell.

The instant invention further embraces an epithelial cell graft. The graft comprises a plurality of the above-described genetically modified cells attached to a support that is suitable for implantation into a mammalian recipient, preferably into the oral cavity. The support can be formed of a natural or synthetic material. According to another aspect of the invention, an encapsulated host cell expression system is provided. The encapsulated system includes a capsule suitable for implantation into a mammalian recipient and a plurality of the above-described genetically modified host cells contained therein. The capsule can be formed of a synthetic or naturally-occurring material. The formulation of such capsules is known to one of ordinary skill in the art. In contrast to the host cells that are directly implanted into the mammalian recipient (*i.e.*, implanted in a manner such that the genetically modified cells are in direct physical contact with the host cell-compatible site), the encapsulated cells remain isolated (*i.e.*, not in direct physical contact with the site) following implantation. Thus, the encapsulated host cell system is not limited to a

capsule including genetically-modified non-immortalized host cells, but may contain genetically modified immortalized host cells.

The invention has been disclosed broadly and illustrated in reference to representative embodiments described above. Those skilled in the art will recognize that various modifications can be made to the present invention without departing from the spirit and scope thereof.

We claim:

1. A purified polypeptide comprising an amino acid sequence having at least 95% sequence identity with sequences selected from the group consisting of SEQ ID NOS: 1, 4, 6, 7, 8 and 9, wherein said polypeptide is capable of regulating signal transduction.

2. The polypeptide of claim 1, wherein said polypeptide is capable of catalyzing the transfer of a phosphate group from a donor molecule to an acceptor molecule.

3. An isolated DNA molecule encoding the polypeptide of claim 1.

4. An isolated DNA molecule comprising the nucleotide sequence selected from the group consisting of SEQ ID NOS: 2, 3 and 5.

5. A vector comprising the DNA molecule of claim 3.

6. A host cell transformed with the vector of claim 5.

7. The host cell of claim 6, wherein said host cell produces a polypeptide capable of regulating signal transduction.

8. The host cell of claim 6, wherein said host cell produces a polypeptide capable of catalyzing the transfer of a phosphate group from a donor molecule to an acceptor molecule.

9. An isolated antibody against the polypeptide of claim 1.

10. The antibody of claim 9, wherein said antibody is a monoclonal antibody.

11. The antibody of claim 9, wherein said antibody is capable of inhibiting the regulation of signal transduction.

12. The antibody of claim 9, wherein said antibody is capable of inhibiting the transfer of a phosphate group from a donor molecule to an acceptor molecule.

13. An isolated nucleic acid capable of hybridizing under high stringency conditions to the DNA molecule of claim 3.

14. The isolated nucleic acid of claim 13, wherein said nucleic acid is capable of inhibiting the regulation of signal transduction.

15. The isolated nucleic acid of claim 13, wherein said nucleic acid is capable of inhibiting said transfer of said phosphate group from said donor molecule to said acceptor molecule.

16. A method of preventing or treating disease in a mammal comprising administering to said mammal an effective amount of material, selected from the group consisting of the polypeptide of claim 1 and antibody against said polypeptide, in a pharmaceutically acceptable sterile vehicle.

17. The method of claim 16, wherein said mammal is a human.

18. The method of claim 16, wherein said disease is cardiac disease.

19. A vaccine for preventing disease in a mammal comprising administering to said mammal an effective amount of material, selected from the group consisting of the polypeptide of claim 1 and the antibody against said polypeptide, in a pharmaceutically acceptable sterile vehicle.

20. The vaccine of claim 19, wherein said mammal is a human.

21. The vaccine of claim 19, wherein said disease is cardiac disease.

22. A method of preventing or treating disease in a mammal comprising administering to said mammal syngeneic cells transformed with the vector of claim 5, wherein said transformed syngeneic cells produce a polypeptide capable of regulating signal transduction.

23. The method of claim 22, wherein said mammal is a human.

24. The method of claim 22, wherein said disease is cardiac disease.

25. A method of preventing or treating disease in a mammal comprising administering to said mammal syngeneic cells transformed with the vector of claim 5, wherein said transformed syngeneic cells produce a polypeptide capable of catalyzing the transfer of a phosphate group from a donor molecule to an acceptor molecule.

26. The method of claim 25, wherein said mammal is a human.

27. The method of claim 25, wherein said disease is cardiac disease.

28. A kit for detecting the expression of a protein capable of regulating signal transduction, comprising the polypeptide of claim 1.

29. The kit of claim 28, further comprising a detectable label selected from the group consisting of colorimetric, enzymatic, chemiluminescent, fluorescent and radioactive labels.

30. A kit for detecting the expression of a protein capable of acting as a donor molecule or an acceptor molecule of a phosphate group, comprising the polypeptide of claim 1.

31. The kit of claim 30, further comprising a detectable label selected from the group consisting of colorimetric, enzymatic, chemiluminescent, fluorescent and radioactive labels.

32. A method for detecting the expression of a protein capable of regulating signal transduction, comprising contacting a sample with the polypeptide of claim 1, and detecting any effect of the sample on an indicator of signal transduction.

33. The method of claim 32, wherein said polypeptide is immobilized to a solid support.

34. The method of claim 32, wherein said phosphate group is detectably labeled.

35. A method for detecting the expression of a protein capable of acting as a donor molecule or an acceptor molecule of a phosphate group, comprising contacting a sample with the polypeptide of claim 1, and detecting any transfer of said phosphate group.

36. The method of claim 35, wherein said polypeptide is immobilized to a solid support.

37. The method of claim 35, wherein said phosphate group is detectably labeled.

SEQUENCE LISTING

<110> Wenlin Zeng
Lawrence Stanton

<120> Mediators of signal transduction

<130> SCIOS.021A

<160> 9

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1351
<212> PRT
<213> Homo sapiens

<400> 1

```
Gly Val Gln Phe Ser Gln Tyr Gly Ser Pro Glu Phe Val Ser Pro Glu
 1               5                  10                  15
Ile Ile Gln Gln Asn Pro Val Ser Glu Ala Ser Asp Ile Trp Ala Met
            20                  25                  30
Gly Val Ile Ser Tyr Leu Ser Leu Thr Cys Ser Ser Pro Phe Ala Gly
        35                  40                  45
Glu Ser Asp Arg Ala Thr Leu Leu Asn Val Leu Glu Gly Arg Val Ser
    50                  55                  60
Trp Ser Ser Pro Met Ala Ala His Leu Ser Glu Asp Ala Lys Asp Phe
65                  70                  75                  80
Ile Lys Ala Thr Leu Gln Arg Ala Pro Gln Ala Arg Pro Ser Ala Ala
                85                  90                  95
Gln Cys Leu Ser His Pro Trp Phe Leu Lys Ser Met Pro Ala Glu Glu
            100                 105                 110
Ala His Phe Ile Asn Thr Lys Gln Leu Lys Phe Leu Leu Ala Arg Ser
        115                 120                 125
Arg Trp Gln Arg Ser Leu Met Ser Tyr Lys Ser Ile Leu Val Met Arg
    130                 135                 140
Ser Ile Pro Glu Leu Leu Arg Gly Pro Pro Asp Ser Pro Ser Leu Gly
145                 150                 155                 160
Val Ala Arg His Leu Cys Arg Asp Thr Gly Gly Ser Ser Ser Ser Ser
                165                 170                 175
Ser Ser Ser Asp Asn Glu Leu Ala Pro Phe Ala Arg Ala Lys Ser Leu
            180                 185                 190
Pro Pro Ser Pro Val Thr His Ser Pro Leu Leu His Pro Arg Gly Phe
        195                 200                 205
Leu Arg Pro Ser Ala Ser Leu Pro Glu Glu Ala Glu Ala Ser Glu Arg
    210                 215                 220
Ser Thr Glu Ala Pro Ala Pro Pro Ala Ser Pro Glu Gly Ala Gly Pro
225                 230                 235                 240
Pro Ala Ala Gln Gly Cys Val Pro Arg His Ser Val Ile Arg Ser Leu
                245                 250                 255
Phe Tyr His Gln Ala Gly Glu Ser Pro Glu His Gly Ala Leu Ala Pro
            260                 265                 270
```

```
Gly Ser Arg Arg His Pro Ala Arg Arg Arg His Leu Leu Lys Gly Gly
        275                 280                 285
Tyr Ile Ala Gly Ala Leu Pro Gly Leu Arg Glu Pro Leu Met Glu His
    290                 295                 300
Arg Val Leu Glu Glu Glu Ala Ala Arg Glu Glu Gln Ala Thr Leu Leu
305                 310                 315                 320
Ala Lys Ala Pro Ser Phe Glu Thr Ala Leu Arg Leu Pro Ala Ser Gly
                325                 330                 335
Thr His Leu Ala Pro Gly His Ser His Ser Leu Glu His Asp Ser Pro
            340                 345                 350
Ser Thr Pro Arg Pro Ser Ser Glu Ala Cys Gly Glu Ala Gln Arg Leu
        355                 360                 365
Pro Ser Ala Pro Ser Gly Gly Ala Pro Ile Arg Asp Met Gly His Pro
    370                 375                 380
Gln Gly Ser Lys Gln Leu Pro Ser Thr Gly Gly His Pro Gly Thr Ala
385                 390                 395                 400
Gln Pro Glu Arg Pro Ser Pro Asp Ser Pro Trp Gly Gln Pro Ala Pro
                405                 410                 415
Phe Cys His Pro Lys Gln Gly Ser Ala Pro Gln Glu Gly Cys Ser Pro
            420                 425                 430
His Pro Ala Val Ala Pro Cys Pro Pro Gly Ser Phe Pro Pro Gly Ser
        435                 440                 445
Cys Lys Glu Ala Pro Leu Val Pro Ser Ser Pro Phe Leu Gly Gln Pro
    450                 455                 460
Gln Ala Pro Pro Ala Pro Ala Lys Ala Ser Pro Pro Leu Asp Ser Lys
465                 470                 475                 480
Met Gly Pro Gly Asp Ile Ser Leu Pro Gly Arg Pro Lys Pro Gly Pro
                485                 490                 495
Cys Ser Ser Pro Gly Ser Ala Ser Gln Ala Ser Ser Ser Gln Val Ser
            500                 505                 510
Ser Leu Arg Val Gly Ser Ser Gln Val Gly Thr Glu Pro Gly Pro Ser
        515                 520                 525
Leu Asp Ala Glu Gly Trp Thr Gln Glu Ala Glu Asp Leu Ser Asp Ser
    530                 535                 540
Thr Pro Thr Leu Gln Arg Pro Gln Glu Gln Ala Thr Met Arg Lys Phe
545                 550                 555                 560
Ser Leu Gly Gly Arg Gly Gly Tyr Ala Gly Val Ala Gly Tyr Gly Thr
                565                 570                 575
Phe Ala Phe Gly Gly Asp Ala Gly Gly Met Leu Gly Gln Gly Pro Met
            580                 585                 590
Trp Ala Arg Ile Ala Trp Ala Val Ser Gln Ser Glu Glu Glu Glu Gln
        595                 600                 605
Glu Glu Ala Arg Ala Glu Ser Gln Ser Glu Glu Gln Gln Glu Ala Arg
    610                 615                 620
Ala Glu Ser Pro Leu Pro Gln Val Ser Ala Arg Pro Val Pro Glu Val
625                 630                 635                 640
Gly Arg Ala Pro Thr Arg Ser Ser Pro Glu Pro Thr Pro Trp Glu Asp
                645                 650                 655
Ile Gly Gln Val Ser Leu Val Gln Ile Arg Asp Leu Ser Gly Asp Ala
            660                 665                 670
Glu Ala Ala Asp Thr Ile Ser Leu Asp Ile Ser Glu Val Asp Pro Ala
        675                 680                 685
Tyr Leu Asn Leu Ser Asp Leu Tyr Asp Ile Lys Tyr Leu Pro Phe Glu
    690                 695                 700
Phe Met Ile Phe Arg Lys Val Pro Lys Ser Ala Gln Pro Glu Pro Pro
705                 710                 715                 720
Ser Pro Met Ala Glu Glu Glu Leu Ala Glu Phe Pro Glu Pro Thr Trp
```

```
                725                 730                 735
Pro Trp Pro Gly Glu Leu Gly Pro His Ala Gly Leu Glu Ile Thr Glu
            740                 745                 750
Glu Ser Glu Asp Val Asp Ala Leu Leu Ala Glu Ala Ala Val Gly Arg
        755                 760                 765
Lys Arg Lys Trp Ser Ser Pro Ser Arg Ser Leu Phe His Phe Pro Gly
    770                 775                 780
Arg His Leu Pro Leu Asp Glu Pro Ala Glu Leu Gly Leu Arg Glu Arg
785                 790                 795                 800
Val Lys Ala Ser Val Glu His Ile Ser Arg Ile Leu Lys Gly Arg Pro
                805                 810                 815
Glu Gly Leu Glu Lys Glu Gly Pro Pro Arg Lys Lys Pro Gly Leu Ala
            820                 825                 830
Ser Phe Arg Leu Ser Gly Leu Lys Ser Trp Asp Arg Ala Pro Thr Phe
        835                 840                 845
Leu Arg Glu Leu Ser Asp Glu Thr Val Val Leu Gly Gln Ser Val Thr
    850                 855                 860
Leu Ala Cys Gln Val Ser Ala Gln Pro Ala Ala Gln Ala Thr Trp Ser
865                 870                 875                 880
Lys Asp Gly Ala Pro Leu Glu Ser Ser Ser Arg Val Leu Ile Ser Ala
                885                 890                 895
Thr Leu Lys Asn Phe Gln Leu Leu Thr Ile Leu Val Val Val Ala Glu
            900                 905                 910
Asp Leu Gly Val Tyr Thr Cys Ser Val Ser Asn Ala Leu Gly Thr Val
        915                 920                 925
Thr Thr Thr Gly Val Leu Arg Lys Ala Glu Arg Pro Ser Ser Ser Pro
    930                 935                 940
Cys Pro Asp Ile Gly Glu Val Tyr Ala Asp Gly Val Leu Leu Val Trp
945                 950                 955                 960
Lys Pro Val Glu Ser Tyr Gly Pro Val Thr Tyr Ile Val Gln Cys Ser
                965                 970                 975
Leu Glu Gly Gly Ser Trp Thr Thr Leu Ala Ser Asp Ile Phe Asp Cys
            980                 985                 990
Cys Tyr Leu Thr Ser Lys Leu Ser Arg Gly Gly Thr Tyr Thr Phe Arg
        995                 1000                1005
Thr Ala Cys Val Ser Lys Ala Gly Met Gly Pro Tyr Ser Ser Pro Ser
    1010                1015                1020
Glu Gln Val Leu Leu Gly Gly Pro Ser His Leu Ala Ser Glu Glu Glu
1025                1030                1035                1040
Ser Gln Gly Arg Ser Ala Gln Pro Leu Pro Ser Thr Lys Thr Phe Ala
                1045                1050                1055
Phe Gln Thr Gln Ile Gln Arg Gly Arg Phe Ser Val Val Arg Gln Cys
            1060                1065                1070
Trp Glu Lys Ala Ser Gly Arg Ala Leu Ala Ala Lys Ile Ile Pro Tyr
        1075                1080                1085
His Pro Lys Asp Lys Thr Ala Val Leu Arg Glu Tyr Glu Ala Leu Lys
    1090                1095                1100
Gly Leu Arg His Pro His Leu Ala Gln Leu His Ala Ala Tyr Leu Ser
1105                1110                1115                1120
Pro Arg His Leu Val Leu Ile Leu Glu Leu Cys Ser Gly Pro Glu Leu
                1125                1130                1135
Leu Pro Cys Leu Ala Glu Arg Ala Ser Tyr Ser Glu Ser Glu Val Lys
            1140                1145                1150
Asp Tyr Leu Trp Gln Met Leu Ser Ala Thr Gln Tyr Leu His Asn Gln
        1155                1160                1165
His Ile Leu His Leu Asp Leu Arg Ser Glu Asn Met Ile Ile Thr Glu
    1170                1175                1180
```

```
Tyr Asn Leu Leu Lys Val Val Asp Leu Gly Asn Ala Gln Ser Leu Ser
1185                1190                1195                1200
Gln Glu Lys Val Leu Pro Ser Asp Lys Phe Lys Asp Tyr Leu Glu Thr
                1205                1210                1215
Met Ala Pro Glu Leu Leu Glu Gly Gln Gly Ala Val Pro Gln Thr Asp
            1220                1225                1230
Ile Trp Ala Ile Gly Val Thr Ala Phe Ile Met Leu Ser Ala Glu Tyr
        1235                1240                1245
Pro Val Ser Ser Glu Gly Ala Arg Asp Leu Gln Arg Gly Leu Arg Lys
    1250                1255                1260
Gly Leu Val Arg Leu Ser Arg Cys Tyr Ala Gly Leu Ser Gly Gly Ala
1265                1270                1275                1280
Val Ala Phe Leu Arg Ser Thr Leu Cys Ala Gln Pro Trp Gly Arg Pro
                1285                1290                1295
Cys Ala Ser Ser Cys Leu Gln Cys Pro Trp Leu Thr Glu Glu Gly Pro
            1300                1305                1310
Ala Cys Ser Arg Pro Ala Pro Val Thr Phe Pro Thr Ala Arg Leu Arg
        1315                1320                1325
Val Phe Val Arg Asn Arg Glu Lys Arg Arg Ala Leu Leu Tyr Lys Arg
    1330                1335                1340
His Asn Leu Ala Gln Val Arg
1345                1350
```

<210> 2
<211> 4175
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)...(4053)

<400> 2

```
gga gtg cag ttc agc cag tac ggc tcc cct gag ttc gtc tcc ccc gag       48
atc atc cag cag aac cct gtg agc gaa gcc tcc gac att tgg gcc atg       96
ggt gtc atc tcc tac ctc agc ctg acc tgc tca tcc cca ttt gcc ggc      144
gag agt gac cgt gcc acc ctc ctg aac gtc ctg gag ggg cgc gtg tca      192
tgg agc agc ccc atg gct gcc cac ctc agc gaa gac gcc aaa gac ttc      240
atc aag gct acg ctg cag aga gcc cct cag gcc cgg cct agt gcg gcc      288
cag tgc ctc tcc cac ccc tgg ttc ctg aaa tcc atg cct gcg gag gag      336
gcc cac ttc atc aac acc aag cag ctc aag ttc ctc ctg gcc cga agt      384
cgc tgg cag cgt tcc ctg atg agc tac aag tcc atc ctg gtg atg cgc      432
tcc atc cct gag ctg ctg cgg ggc cca ccc gac agc ccc tcc ctc ggc      480
gta gcc cgg cac ctc tgc agg gac act ggt ggc tcc tcc agt tcc tcc      528
tcc tcc tct gac aac gag ctc gcc cca ttt gcc cgg gct aag tca ctg      576
cca ccc tcc ccg gtg aca cac tca cca ctg ctg cac ccc cgg ggc ttc      624
ctg cgg ccc tcg gcc agc ctg cct gag gaa gcc gag gcc agt gag cgc      672
tcc acc gag gcc cca gct ccg cct gca tct ccc gag ggt gcc ggg cca      720
ccg gcc gcc cag ggc tgc gtg ccc cgg cac agc gtc atc cgc agc ctg      768
ttc tac cac cag gcg ggt gag agc cct gag cac ggg gcc ctg gcc ccg      816
ggg agc agg cgg cac ccg gcc cgg cgg cgg cac ctg ctg aag ggc ggg      864
tac att gcg ggg gcg ctg cca ggc ctg cgc gag cca ctg atg gag cac      912
cgc gtg ctg gag gag gag gcc gcc agg gag gag cag gcc acc ctc ctg      960
```

```
gcc aaa gcc ccc tca ttc gag act gcc ctc cgg ctg cct gcc tct ggc     1008
acc cac ttg gcc cct ggc cac agc cac tcc ctg gaa cat gac tct ccg     1056
agc acc ccc cgc ccc tcc tcg gag gcc tgc ggt gag gca cag cga ctg     1104
cct tca gcc ccc tcc ggg ggg gcc cct atc agg gac atg ggg cac cct     1152
cag ggc tcc aag cag ctt cca tcc act ggt ggc cac cca ggc act gct     1200
cag cca gag agg cca tcc ccg gac agc cct tgg ggg cag cca gcc cct     1248
ttc tgc cac ccc aag cag ggt tct gcc ccc cag gag ggc tgc agc ccc     1296
cac cca gca gtt gcc cca tgc cct cct ggc tcc ttc cct cca gga tct     1344
tgc aaa gag gcc ccc tta gta ccc tca agc ccc ttc ttg gga cag ccc     1392
cag gca ccc cct gcc cct gcc aaa gca agc ccc cca ttg gac tct aag     1440
atg ggg cct gga gac atc tct ctt cct ggg agg cca aaa ccc ggc ccc     1488
tgc agt tcc cca ggg tca gcc tcc cag gcg agc tct tcc caa gtg agc     1536
tcc ctc agg gtg ggc tcc tcc cag gtg ggc aca gag cct ggc ccc tcc     1584
ctg gat gcg gag ggc tgg acc cag gag gct gag gat ctg tcc gac tcc     1632
aca ccc acc ttg cag cgg cct cag gaa cag gcg acc atg cgc aag ttc     1680
tcc ctg ggt ggt cgc ggg ggc tac gca ggc gtg gct ggc tat ggc acc     1728
ttt gcc ttt ggt gga gat gca ggg ggc atg ctg ggg cag ggg ccc atg     1776
tgg gcc agg ata gcc tgg gct gtg tcc cag tca gag gag gag gag gag cag     1824
gag gag gcc agg gct gag tcc cag tcg gag gag cag cag gag gcc agg     1872
gct gag agc cca ctg ccc cag gtc agt gca agg cct gtg cct gag gtc     1920
ggc agg gct ccc acc agg agc tct cca gag ccc acc cca tgg gag gac     1968
atc ggg cag gtc tcc ctg gtg cag atc cgg gac ctg tca ggt gat gcg     2016
gag gcg gcc gac aca ata tcc ctg gac att tcc gag gtg gac ccc gcc     2064
tac ctc aac ctc tca gac ctg tac gat atc aag tac ctc cca ttc gag     2112
ttt atg atc ttc agg aaa gtc ccc aag tcc gct cag cca gag ccg ccc     2160
tcc ccc atg gct gag gag gag ctg gcc gag ttc ccg gag ccc acg tgg     2208
ccc tgg cca ggt gaa ctg ggc ccc cac gca ggc ctg gag atc aca gag     2256
gag tca gag gat gtg gac gcg ctg ctg gca gag gct gcc gtg ggc agg     2304
aag cgc aag tgg tcc tcg ccg tca cgc agc ctc ttc cac ttc cct ggg     2352
agg cac ctg ccg ctg gat gag cct gca gag ctg ggg ctg cgt gag aga     2400
gtg aag gcc tcc gtg gag cac atc tcc cgg atc ctg aag ggc agg ccg     2448
gaa ggt ctg gag aag gag ggg ccc ccc agg aag aag cca ggc ctt gct     2496
tcc ttc cgg ctc tca ggt ctg aag agc tgg gac cga gcg ccg aca ttc     2544
cta agg gag ctc tca gat gag act gtg gtc ctg ggc cag tca gtg aca     2592
ctg gcc tgc cag gtg tca gcc cag cca gct gcc cag gcc acc tgg agc     2640
aaa gac gga gcc ccc ctg gag agc agc agc cgt gtc ctc atc tct gcc     2688
acc ctc aag aac ttc cag ctt ctg acc atc ctg gtg gtg gtg gct gag     2736
gac ctg ggt gtg tac acc tgc agc gtg agc aat gcg ctg ggg aca gtg     2784
acc acc acg ggc gtc ctc cgg aag gca gag cgc ccc tca tct tcg cca     2832
tgc ccg gat atc ggg gag gtg tac gcg gat ggg gtg ctg ctg gtc tgg     2880
aag ccc gtg gaa tcc tac ggc cct gtg acc tac att gtg cag tgc agc     2928
cta gaa ggc ggc agc tgg acc aca ctg gcc tcc gac atc ttt gac tgc     2976
tgc tac ctg acc agc aag ctc tcc cgg ggt ggc acc tac acc ttc cgc     3024
acg gca tgt gtc agc aag gca gga atg ggt ccc tac agc agc ccc tcg     3072
gag caa gtc ctc ctg gga ggg ccc agc cac ctg gcc tct gag gag gag     3120
agc cag ggg cgg tca gcc caa ccc ctg ccc agc aca aag acc ttc gca     3168
ttc cag aca cag atc cag agg ggc cgc ttc agc gtg gtg cgg caa tgc     3216
tgg gag aag gcc agc ggg cgg gcg ctg gcc gcc aag atc atc ccc tac     3264
cac ccc aag gac aag aca gca gtg ctg cgc gaa tac gag gcc ctc aag     3312
ggc ctg cgc cac ccg cac ctg gcc cag ctg cac gca gcc tac ctc agc     3360
ccc cgg cac ctg gtg ctc atc ttg gag ctg tgc tct ggg ccc gag ctg     3408
ctc ccc tgc ctg gcc gag agg gcc tcc tac tca gaa tcc gag gtg aag     3456
gac tac ctg tgg cag atg ttg agt gcc acc cag tac ctg cac aac cag     3504
cac atc ctg cac ctg gac ctg agg tcc gag aac atg atc atc acc gaa     3552
tac aac ctg ctc aag gtc gtg gac ctg ggc aat gca cag agc ctc agc     3600
cag gag aag gtg ctg ccc tca gac aag ttc aag gac tac cta gag acc     3648
atg gct cca gag ctc ctg gag ggc cag ggg gct gtt cca cag aca gac     3696
```

atc tgg gcc atc ggt gtg aca gcc ttc atc atg ctg agc gcc gag tac 3744
ccg gtg agc agc gag ggt gca cgc gac ctg cag aga gga ctg cgc aag 3792
ggg ctg gtc cgg ctg agc cgc tgc tac gcg ggg ctg tcc ggg ggc gcc 3840
gtg gcc ttc ctg cgc agc act ctg tgc gcc cag ccc tgg ggc cgg ccc 3888
tgc gcg tcc agc tgc ctg cag tgc ccg tgg cta aca gag gag ggc ccg 3936
gcc tgt tcg cgg ccc gcg ccc gtg acc ttc cct acc gcg cgg ctg cgc 3984
gtc ttc gtg cgc aat cgc gag aag aga cgc gcg ctg ctg tac aag agg 4032
cac aac ctg gcc cag gtg cgc tgagggtcgc cccggccaca cccttggtct 4083
ccccgctggg ggtcgctgca gacgcgccaa taaaaacgca cagccgggcg agaaaaaaaa 4143
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 4175

<210> 3
<211> 5007
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (97)...(4926)

<400> 3
ctgctctggg acgtccacag ccacgtggtc agagagacca cacagaggac ctacacatac 60
caggccatcg acacgcacac cgcacggccc ccatcc atg cag gta acc atc gag 114
gat gtg cag gca cag aca ggc gga acg gcc caa ttc gag gct atc att 162
gag ggc gac cca cag ccc tcg gtg acc tgg tac aag gac agc gtc cag 210
ctg gtg gac agc acc cgg ctt agc cag cag caa gaa ggc acc aca tac 258
tcc ctg gtg ctg agg cat gtg gcc tcg aag gat gcc ggc gtt tac acc 306
tgc ctg gcc caa aac act ggt ggc cag gtg ctc tgc aag gca gag ctg 354
ctg gtg ctt ggg ggg gac aat gag ccg gac tca gag aag caa agc cac 402
cgg agg aag ctg cac tcc ttc tat gag gtc aag gag gag att gga agg 450
ggc gtg ttt ggc ttc gta aaa aga gtg cag cac aaa gga aac aag atc 498
ttg tgc gct gcc aag ttc atc ccc cta cgg agc aga act cgg gcc cag 546
gca tac agg gag cga gac atc ctg gcc gcg ctg agc cac ccg ctg gtc 594
acg ggg ctg ctg gac cag ttt gag acc cgc aag acc ctc atc ctc atc 642
ctg gag ctg tgc tca tcc gag gag ctg ctg gac cgc ctg tac agg aag 690
ggc gtg gtg acg gag gcc gag gtc aag gtc tac atc cag cag ctg gtg 738
gag ggg ctg cac tac ctg cac agc cat ggc gtt ctc cac ctg gac ata 786
aag ccc tct aac atc ctg atg gtg cat cct gcc cgg gaa gac att aaa 834
atc tgc gac ttt ggc ttt gcc cag aac atc acc cca gca gag ctg cag 882
ttc agc cag tac ggc tcc cct gag ttc gtc tcc ccc gag atc atc cag 930
cag aac cct gtg agc gaa gcc tcc gac att tgg gcc atg ggt gtc atc 978
tcc tac ctc agc ctg acc tgc tca tcc cca ttt gcc ggc gag agt gac 1026
cgt gcc acc ctc ctg aac gtc ctg gag ggg cgc gtg tca tgg agc agc 1074
ccc atg gct gcc cac ctc agc gaa gac gcc aaa gac ttc atc aag gct 1122
acg ctg cag aga gcc cct cag gcc cgg cct agt gcg gcc cag tgc ctc 1170
tcc cac ccc tgg ttc ctg aaa tcc atg cct gcg gag gag gcc cac ttc 1218
atc aac acc aag cag ctc aag ttc ctc ctg gcc cga agt cgc tgg cag 1266
cgt tcc ctg atg agc tac aag tcc atc ctg gtg atg cgc tcc atc cct 1314
gag ctg ctg cgg ggc cca ccc gac agc ccc tcc ctc ggc gta gcc cgg 1362
cac ctc tgc agg gac act ggt ggc tcc tcc agt tcc tcc tcc tcc tct 1410
gac aac gag ctc gcc cca ttt gcc cgg gct aag tca ctg cca ccc tcc 1458
ccg gtg aca cac tca cca ctg ctg cac ccc cgg ggc ttc ctg cgg ccc 1506

```
tcg gcc agc ctg cct gag gaa gcc gag gcc agt gag cgc tcc acc gag     1554
gcc cca gct ccg cct gca tct ccc gag ggt gcc ggg cca ccg gcc gcc     1602
cag ggc tgc gtg ccc cgg cac agc gtc atc cgc agc ctg ttc tac cac     1650
cag gcg ggt gag agc cct gag cac ggg gcc ctg gcc ccg ggg agc agg     1698
cgg cac ccg gcc cgg cgg cgg cac ctg ctg aag ggc ggg tac att gcg     1746
ggg gcg ctg cca ggc ctg cgc gag cca ctg atg gag cac cgc gtg ctg     1794
gag gag gag gcc gcc agg gag gag cag gcc acc ctc ctg gcc aaa gcc     1842
ccc tca ttc gag act gcc ctc cgg ctg cct gcc tct ggc acc cac ttg     1890
gcc cct ggc cac agc cac tcc ctg gaa cat gac tct ccg agc acc ccc     1938
cgc ccc tcc tcg gag gcc tgc ggt gag gca cag cga ctg cct tca gcc     1986
ccc tcc ggg ggg gcc cct atc agg gac atg ggg cac cct cag ggc tcc     2034
aag cag ctt cca tcc act ggt ggc cac cca ggc act gct cag cca gag     2082
agg cca tcc ccg gac agc cct tgg ggg cag cca gcc cct ttc tgc cac     2130
ccc aag cag ggt tct gcc ccc cag gag ggc tgc agc ccc cac cca gca     2178
gtt gcc cca tgc cct cct ggc tcc ttc cct cca gga tct tgc aaa gag     2226
gcc ccc tta gta ccc tca agc ccc ttc ttg gga cag ccc cag gca ccc     2274
cct gcc cct gcc aaa gca agc ccc cca ttg gac tct aag atg ggg cct     2322
gga gac atc tct ctt cct ggg agg cca aaa ccc ggc ccc tgc agt tcc     2370
cca ggg tca gcc tcc cag gcg agc tct tcc caa gtg agc tcc ctc agg     2418
gtg ggc tcc tcc cag gtg ggc aca gag cct ggc ccc tcc ctg gat gcg     2466
gag ggc tgg acc cag gag gct gag gat ctg tcc gac tcc aca ccc acc     2514
ttg cag cgg cct cag gaa cag gcg acc atg cgc aag ttc tcc ctg ggt     2562
ggt cgc ggg ggc tac gca ggc gtg gct ggc tat ggc acc ttt gcc ttt     2610
ggt gga gat gca ggg ggc atg ctg ggg cag ggg ccc atg tgg gcc agg     2658
ata gcc tgg gct gtg tcc cag tca gag gag gag gag cag gag gag gcc     2706
agg gct gag tcc cag tcg gag gag cag cag gag gcc agg gct gag agc     2754
cca ctg ccc cag gtc agt gca agg cct gtg cct gag gtc ggc agg gct     2802
ccc acc agg agc tct cca gag ccc acc cca tgg gag gac atc ggg cag     2850
gtc tcc ctg gtg cag atc cgg gac ctg tca ggt gat gcg gag gcg gcc     2898
gac aca ata tcc ctg gac att tcc gag gtg gac ccc gcc tac ctc aac     2946
ctc tca gac ctg tac gat atc aag tac ctc cca ttc gag ttt atg atc     2994
ttc agg aaa gtc ccc aag tcc gct cag cca gag ccg ccc tcc ccc atg     3042
gct gag gag gag ctg gcc gag ttc ccg gag ccc acg tgg ccc tgg cca     3090
ggt gaa ctg ggc ccc cac gca ggc ctg gag atc aca gag gag tca gag     3138
gat gtg gac gcg ctg ctg gca gag gct gcc gtg ggc agg aag cgc aag     3186
tgg tcc tcg ccg tca cgc agc ctc ttc cac ttc cct ggg agg cac ctg     3234
ccg ctg gat gag cct gca gag ctg ggg ctg cgt gag aga gtg aag gcc     3282
tcc gtg gag cac atc tcc cgg atc ctg aag ggc agg ccg gaa ggt ctg     3330
gag aag gag ggg ccc ccc agg aag aag cca ggc ctt gct tcc ttc cgg     3378
ctc tca ggt ctg aag agc tgg gac cga gcg ccg aca ttc cta agg gag     3426
ctc tca gat gag act gtg gtc ctg ggc cag tca gtg aca ctg gcc tgc     3474
cag gtg tca gcc cag cca gct gcc cag gcc acc tgg agc aaa gac gga     3522
gcc ccc ctg gag agc agc agc cgt gtc ctc atc tct gcc acc ctc aag     3570
aac ttc cag ctt ctg acc atc ctg gtg gtg gtg gct gag gac ctg ggt     3618
gtg tac acc tgc agc gtg agc aat gcg ctg ggg aca gtg acc acc acg     3666
ggc gtc ctc cgg aag gca gag cgc ccc tca tct tcg cca tgc ccg gat     3714
atc ggg gag gtg tac gcg gat ggg gtg ctg ctg gtc tgg aag ccc gtg     3762
gaa tcc tac ggc cct gtg acc tac att gtg cag tgc agc cta gaa ggc     3810
ggc agc tgg acc aca ctg gcc tcc gac atc ttt gac tgc tgc tac ctg     3858
acc agc aag ctc tcc cgg ggt ggc acc tac acc ttc cgc acg gca tgt     3906
gtc agc aag gca gga atg ggt ccc tac agc agc ccc tcg gag caa gtc     3954
ctc ctg gga gcg ccc agc cac ctg gcc tct gag gag gag agc cag ggg     4002
cgg tca gcc caa ccc ctg ccc agc aca aag acc ttc gca ttc cag aca     4050
cag atc cag agg ggc cgc ttc agc gtg gtg cgg caa tgc tgg gag aag     4098
gcc agc ggg cgg gcg ctg gcc gcc aag atc atc ccc tac cac ccc aag     4146
gac aag aca gca gtg ctg cgc gaa tac gag gcc ctc aag ggc ctg cgc     4194
cac ccg cac ctg gcc cag ctg cac gca gcc tac ctc agc ccc cgg cac     4242
```

```
ctg gtg ctc atc ttg gag ctg tgc tct ggg ccc gag ctg ctc ccc tgc     4290
ctg gcc gag agg gcc tcc tac tca gaa tcc gag gtg aag gac tac ctg     4338
tgg cag atg ttg agt gcc acc cag tac ctg cac aac cag cac atc ctg     4386
cac ctg gac ctg agg tcc gag aac atg atc atc acc gaa tac aac ctg     4434
ctc aag gtc gtg gac ctg ggc aat gca cag agc ctc agc cag gag aag     4482
gtg ctg ccc tca gac aag ttc aag gac tac cta gag acc atg gct cca     4530
gag ctc ctg gag ggc cag ggg gct gtt cca cag aca gac atc tgg gcc     4578
atc ggt gtg aca gcc ttc atc atg ctg agc gcc gag tac ccg gtg agc     4626
agc gag ggt gca cgc gac ctg cag aga gga ctg cgc aag ggg ctg gtc     4674
cgg ctg agc cgc tgc tac gcg ggg ctg tcc ggg ggc gcc gtg gcc ttc     4722
ctg cgc agc act ctg tgc gcc cag ccc tgg ggc cgg ccc tgc gcg tcc     4770
agc tgc ctg cag tgc ccg tgg cta aca gag gag ggc ccg gcc tgt tcg     4818
cgg ccc gcg ccc gtg acc ttc cct acc gcg cgg ctg cgc gtc ttc gtg     4866
cgc aat cgc gag aag aga cgc gcg ctg ctg tac aag agg cac aac ctg     4914
gcc cag gtg cgc tgagggtcgc cccggccaca cccttggtct ccccgctggg        4966
ggtcgctgca gacgcgccaa taaaaacgca cagccgggcg a                     5007
```

<210> 4
<211> 1610
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Gln Val Thr Ile Glu Asp Val Gln Ala Gln Thr Gly Gly Thr Ala
 1               5                  10                  15
Gln Phe Glu Ala Ile Ile Glu Gly Asp Pro Gln Pro Ser Val Thr Trp
            20                  25                  30
Tyr Lys Asp Ser Val Gln Leu Val Asp Ser Thr Arg Leu Ser Gln Gln
        35                  40                  45
Gln Glu Gly Thr Thr Tyr Ser Leu Val Leu Arg His Val Ala Ser Lys
    50                  55                  60
Asp Ala Gly Val Tyr Thr Cys Leu Ala Gln Asn Thr Gly Gly Gln Val
65                  70                  75                  80
Leu Cys Lys Ala Glu Leu Leu Val Leu Gly Gly Asp Asn Glu Pro Asp
                85                  90                  95
Ser Glu Lys Gln Ser His Arg Arg Lys Leu His Ser Phe Tyr Glu Val
            100                 105                 110
Lys Glu Glu Ile Gly Arg Gly Val Phe Gly Phe Val Lys Arg Val Gln
        115                 120                 125
His Lys Gly Asn Lys Ile Leu Cys Ala Ala Lys Phe Ile Pro Leu Arg
    130                 135                 140
Ser Arg Thr Arg Ala Gln Ala Tyr Arg Glu Arg Asp Ile Leu Ala Ala
145                 150                 155                 160
Leu Ser His Pro Leu Val Thr Gly Leu Leu Asp Gln Phe Glu Thr Arg
                165                 170                 175
Lys Thr Leu Ile Leu Ile Leu Glu Leu Cys Ser Ser Glu Glu Leu Leu
            180                 185                 190
Asp Arg Leu Tyr Arg Lys Gly Val Val Thr Glu Ala Glu Val Lys Val
        195                 200                 205
Tyr Ile Gln Gln Leu Val Glu Gly Leu His Tyr Leu His Ser His Gly
    210                 215                 220
Val Leu His Leu Asp Ile Lys Pro Ser Asn Ile Leu Met Val His Pro
```

```
225                 230                 235                 240
Ala Arg Glu Asp Ile Lys Ile Cys Asp Phe Gly Phe Ala Gln Asn Ile
                245                 250                 255
Thr Pro Ala Glu Leu Gln Phe Ser Gln Tyr Gly Ser Pro Glu Phe Val
            260                 265                 270
Ser Pro Glu Ile Ile Gln Gln Asn Pro Val Ser Glu Ala Ser Asp Ile
        275                 280                 285
Trp Ala Met Gly Val Ile Ser Tyr Leu Ser Leu Thr Cys Ser Ser Pro
    290                 295                 300
Phe Ala Gly Glu Ser Asp Arg Ala Thr Leu Leu Asn Val Leu Glu Gly
305                 310                 315                 320
Arg Val Ser Trp Ser Ser Pro Met Ala Ala His Leu Ser Glu Asp Ala
                325                 330                 335
Lys Asp Phe Ile Lys Ala Thr Leu Gln Arg Ala Pro Gln Ala Arg Pro
            340                 345                 350
Ser Ala Ala Gln Cys Leu Ser His Pro Trp Phe Leu Lys Ser Met Pro
        355                 360                 365
Ala Glu Glu Ala His Phe Ile Asn Thr Lys Gln Leu Lys Phe Leu Leu
    370                 375                 380
Ala Arg Ser Arg Trp Gln Arg Ser Leu Met Ser Tyr Lys Ser Ile Leu
385                 390                 395                 400
Val Met Arg Ser Ile Pro Glu Leu Leu Arg Gly Pro Pro Asp Ser Pro
                405                 410                 415
Ser Leu Gly Val Ala Arg His Leu Cys Arg Asp Thr Gly Gly Ser Ser
            420                 425                 430
Ser Ser Ser Ser Ser Ser Asp Asn Glu Leu Ala Pro Phe Ala Arg Ala
        435                 440                 445
Lys Ser Leu Pro Pro Ser Pro Val Thr His Ser Pro Leu Leu His Pro
    450                 455                 460
Arg Gly Phe Leu Arg Pro Ser Ala Ser Leu Pro Glu Glu Ala Glu Ala
465                 470                 475                 480
Ser Glu Arg Ser Thr Glu Ala Pro Ala Pro Pro Ala Ser Pro Glu Gly
                485                 490                 495
Ala Gly Pro Pro Ala Ala Gln Gly Cys Val Pro Arg His Ser Val Ile
            500                 505                 510
Arg Ser Leu Phe Tyr His Gln Ala Gly Glu Ser Pro Glu His Gly Ala
        515                 520                 525
Leu Ala Pro Gly Ser Arg Arg His Pro Ala Arg Arg Arg His Leu Leu
    530                 535                 540
Lys Gly Gly Tyr Ile Ala Gly Ala Leu Pro Gly Leu Arg Glu Pro Leu
545                 550                 555                 560
Met Glu His Arg Val Leu Glu Glu Glu Ala Ala Arg Glu Glu Gln Ala
                565                 570                 575
Thr Leu Leu Ala Lys Ala Pro Ser Phe Glu Thr Ala Leu Arg Leu Pro
            580                 585                 590
Ala Ser Gly Thr His Leu Ala Pro Gly His Ser His Ser Leu Glu His
        595                 600                 605
Asp Ser Pro Ser Thr Pro Arg Pro Ser Ser Glu Ala Cys Gly Glu Ala
    610                 615                 620
Gln Arg Leu Pro Ser Ala Pro Ser Gly Gly Ala Pro Ile Arg Asp Met
625                 630                 635                 640
Gly His Pro Gln Gly Ser Lys Gln Leu Pro Ser Thr Gly Gly His Pro
                645                 650                 655
Gly Thr Ala Gln Pro Glu Arg Pro Ser Pro Asp Ser Pro Trp Gly Gln
            660                 665                 670
Pro Ala Pro Phe Cys His Pro Lys Gln Gly Ser Ala Pro Gln Glu Gly
        675                 680                 685
```

```
Cys Ser Pro His Pro Ala Val Ala Pro Cys Pro Pro Gly Ser Phe Pro
    690                 695                 700
Pro Gly Ser Cys Lys Glu Ala Pro Leu Val Pro Ser Ser Pro Phe Leu
705                 710                 715                 720
Gly Gln Pro Gln Ala Pro Pro Ala Pro Ala Lys Ala Ser Pro Pro Leu
                725                 730                 735
Asp Ser Lys Met Gly Pro Gly Asp Ile Ser Leu Pro Gly Arg Pro Lys
            740                 745                 750
Pro Gly Pro Cys Ser Ser Pro Gly Ser Ala Ser Gln Ala Ser Ser Ser
        755                 760                 765
Gln Val Ser Ser Leu Arg Val Gly Ser Ser Gln Val Gly Thr Glu Pro
    770                 775                 780
Gly Pro Ser Leu Asp Ala Glu Gly Trp Thr Gln Glu Ala Glu Asp Leu
785                 790                 795                 800
Ser Asp Ser Thr Pro Thr Leu Gln Arg Pro Gln Glu Gln Ala Thr Met
                805                 810                 815
Arg Lys Phe Ser Leu Gly Gly Arg Gly Gly Tyr Ala Gly Val Ala Gly
            820                 825                 830
Tyr Gly Thr Phe Ala Phe Gly Gly Asp Ala Gly Gly Met Leu Gly Gln
        835                 840                 845
Gly Pro Met Trp Ala Arg Ile Ala Trp Ala Val Ser Gln Ser Glu Glu
    850                 855                 860
Glu Glu Gln Glu Glu Ala Arg Ala Glu Ser Gln Ser Glu Glu Gln Gln
865                 870                 875                 880
Glu Ala Arg Ala Glu Ser Pro Leu Pro Gln Val Ser Ala Arg Pro Val
                885                 890                 895
Pro Glu Val Gly Arg Ala Pro Thr Arg Ser Ser Pro Glu Pro Thr Pro
            900                 905                 910
Trp Glu Asp Ile Gly Gln Val Ser Leu Val Gln Ile Arg Asp Leu Ser
        915                 920                 925
Gly Asp Ala Glu Ala Ala Asp Thr Ile Ser Leu Asp Ile Ser Glu Val
    930                 935                 940
Asp Pro Ala Tyr Leu Asn Leu Ser Asp Leu Tyr Asp Ile Lys Tyr Leu
945                 950                 955                 960
Pro Phe Glu Phe Met Ile Phe Arg Lys Val Pro Lys Ser Ala Gln Pro
                965                 970                 975
Glu Pro Pro Ser Pro Met Ala Glu Glu Glu Leu Ala Glu Phe Pro Glu
            980                 985                 990
Pro Thr Trp Pro Trp Pro Gly Glu Leu Gly Pro His Ala Gly Leu Glu
        995                 1000                1005
Ile Thr Glu Glu Ser Glu Asp Val Asp Ala Leu Leu Ala Glu Ala Ala
    1010                1015                1020
Val Gly Arg Lys Arg Lys Trp Ser Ser Pro Ser Arg Ser Leu Phe His
1025                1030                1035                1040
Phe Pro Gly Arg His Leu Pro Leu Asp Glu Pro Ala Glu Leu Gly Leu
                1045                1050                1055
Arg Glu Arg Val Lys Ala Ser Val Glu His Ile Ser Arg Ile Leu Lys
            1060                1065                1070
Gly Arg Pro Glu Gly Leu Glu Lys Glu Gly Pro Pro Arg Lys Lys Pro
        1075                1080                1085
Gly Leu Ala Ser Phe Arg Leu Ser Gly Leu Lys Ser Trp Asp Arg Ala
    1090                1095                1100
Pro Thr Phe Leu Arg Glu Leu Ser Asp Glu Thr Val Val Leu Gly Gln
1105                1110                1115                1120
Ser Val Thr Leu Ala Cys Gln Val Ser Ala Gln Pro Ala Ala Gln Ala
                1125                1130                1135
Thr Trp Ser Lys Asp Gly Ala Pro Leu Glu Ser Ser Ser Arg Val Leu
```

```
            1140                1145                1150
Ile Ser Ala Thr Leu Lys Asn Phe Gln Leu Leu Thr Ile Leu Val Val
        1155                1160                1165
Val Ala Glu Asp Leu Gly Val Tyr Thr Cys Ser Val Ser Asn Ala Leu
    1170                1175                1180
Gly Thr Val Thr Thr Thr Gly Val Leu Arg Lys Ala Glu Arg Pro Ser
1185                1190                1195                1200
Ser Ser Pro Cys Pro Asp Ile Gly Glu Val Tyr Ala Asp Gly Val Leu
                1205                1210                1215
Leu Val Trp Lys Pro Val Glu Ser Tyr Gly Pro Val Thr Tyr Ile Val
            1220                1225                1230
Gln Cys Ser Leu Glu Gly Gly Ser Trp Thr Thr Leu Ala Ser Asp Ile
        1235                1240                1245
Phe Asp Cys Cys Tyr Leu Thr Ser Lys Leu Ser Arg Gly Gly Thr Tyr
    1250                1255                1260
Thr Phe Arg Thr Ala Cys Val Ser Lys Ala Gly Met Gly Pro Tyr Ser
1265                1270                1275                1280
Ser Pro Ser Glu Gln Val Leu Leu Gly Ala Pro Ser His Leu Ala Ser
                1285                1290                1295
Glu Glu Glu Ser Gln Gly Arg Ser Ala Gln Pro Leu Pro Ser Thr Lys
            1300                1305                1310
Thr Phe Ala Phe Gln Thr Gln Ile Gln Arg Gly Arg Phe Ser Val Val
        1315                1320                1325
Arg Gln Cys Trp Glu Lys Ala Ser Gly Arg Ala Leu Ala Ala Lys Ile
    1330                1335                1340
Ile Pro Tyr His Pro Lys Asp Lys Thr Ala Val Leu Arg Glu Tyr Glu
1345                1350                1355                1360
Ala Leu Lys Gly Leu Arg His Pro His Leu Ala Gln Leu His Ala Ala
                1365                1370                1375
Tyr Leu Ser Pro Arg His Leu Val Leu Ile Leu Glu Leu Cys Ser Gly
            1380                1385                1390
Pro Glu Leu Leu Pro Cys Leu Ala Glu Arg Ala Ser Tyr Ser Glu Ser
        1395                1400                1405
Glu Val Lys Asp Tyr Leu Trp Gln Met Leu Ser Ala Thr Gln Tyr Leu
    1410                1415                1420
His Asn Gln His Ile Leu His Leu Asp Leu Arg Ser Glu Asn Met Ile
1425                1430                1435                1440
Ile Thr Glu Tyr Asn Leu Leu Lys Val Val Asp Leu Gly Asn Ala Gln
                1445                1450                1455
Ser Leu Ser Gln Glu Lys Val Leu Pro Ser Asp Lys Phe Lys Asp Tyr
            1460                1465                1470
Leu Glu Thr Met Ala Pro Glu Leu Leu Glu Gly Gln Gly Ala Val Pro
        1475                1480                1485
Gln Thr Asp Ile Trp Ala Ile Gly Val Thr Ala Phe Ile Met Leu Ser
    1490                1495                1500
Ala Glu Tyr Pro Val Ser Ser Glu Gly Ala Arg Asp Leu Gln Arg Gly
1505                1510                1515                1520
Leu Arg Lys Gly Leu Val Arg Leu Ser Arg Cys Tyr Ala Gly Leu Ser
                1525                1530                1535
Gly Gly Ala Val Ala Phe Leu Arg Ser Thr Leu Cys Ala Gln Pro Trp
            1540                1545                1550
Gly Arg Pro Cys Ala Ser Ser Cys Leu Gln Cys Pro Trp Leu Thr Glu
        1555                1560                1565
Glu Gly Pro Ala Cys Ser Arg Pro Ala Pro Val Thr Phe Pro Thr Ala
    1570                1575                1580
Arg Leu Arg Val Phe Val Arg Asn Arg Glu Lys Arg Arg Ala Leu Leu
1585                1590                1595                1600
```

Tyr Lys Arg His Asn Leu Ala Gln Val Arg
1605 1610

<210> 5
<211> 7928
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (60)...(7847)

<400> 5

```
gaattcggcg cgccagatat cacacgtgcc aaggggctgg ctcactggtg ccgccccga      59
atg ctg gag agg ttc acc ccc aag aaa gtg aag aaa ggc tcc agc atc      107
acc ttc tct gtg aag gta gaa gga cgc ccg gtg ccc acc gtg cac tgg      155
ctc agg gag gag gct gag aga ggc gtg ctg tgg att ggc cct gac aca      203
ccg ggc tac acc gtg gcc agc tct gcg cag cag cac agc ctg gtc ctg      251
ctg gac gtg ggc cgg cag cac cag ggc acc tac aca tgc att gcc agc      299
aac gct gcc ggc cag gcc ctc tgc tcc gcc agc ctg cac gtc tcg ggc      347
ctg cct aag gtg gag gag cag gag aaa gtg aag gaa gcg ctg att tcc      395
act ttc ctg cag ggg acc aca caa gcc atc tca gca cag ggg ttg gaa      443
act gcg agt ttt gct gac ctt ggt ggg cag agg aaa gaa gag cct ctg      491
gct gcc aag gag gcc ctc ggc cac ctg tcc ctc gct gag gtg ggc aca      539
gag gag ttc ctg cag aaa ctg acc tcc cag atc act gag atg gta tcg      587
gcc aag atc acg cag gcc aag ctg cag gtg cca gga ggt gac agt gat      635
gag gac tcc aag aca cca tct gca tcc ccc cgc cat ggc cga tca cgg      683
cca tcc tcc agc atc cag gag tct tcc tca gag tca gag gac ggc gat      731
gcc cga ggc gag atc ttt gac atc tac gtg gtc acc gct gac tac ctg      779
ccc cta ggg gct gag cag gat gcc atc acg ctg cgg gaa ggc cag tat      827
gtg gag gtc ctg gat gca gcc cac cca ctg cgc tgg ctt gtc cgc acc      875
aag ccc acc aag tcc agc ccc tca cgg cag ggc tgg gtg tca cca gcc      923
tac ctg gac agg agg ctc aag ctg tca cct gag tgg ggg gcc gct gag      971
gcc cct gag ttc cct ggg gag gct gtg tct gaa gac gaa tac aag gca     1019
agg ctg agc tct gtg atc cag gag ctg ctg agt tct gag cag gcc ttc     1067
gtg gag gag ctg cag ttc ctg cag agc cac cac ctg cag cac ctg gag     1115
cgc tgc ccc cac gtg ccc ata gcc gtg gcc ggc cag aag gca gtc atc     1163
ttc cgc aat gtg cgg gac atc ggc cgc ttc cac agc agc ttc ctg cag     1211
gag ttg cag cag tgc gac acg gac gac gac gtg gcc atg tgc ttc atc     1259
aag aac cag gcg gcc ttt gag cag tac ctg gag ttc ctg gtg ggg cgt     1307
gtg cag gct gag tcg gtg gtc gtc agc acg gcc atc cag gag ttc tac     1355
aag aaa tac gcg gag gag gcc ctg ttg gca ggg gac ccc tct cag ccc     1403
ccg cca cca cct ctg cag cac tac ctg gag cag cca gtg gag cgg gtg     1451
cag cgc tac cag gcc ttg ctg aag gag ttg atc cgc aac aag gcg cgg     1499
aac aga cag aac tgc gcg ctg ctg gag cag gcc tat gcc gtg gtg tct     1547
gcc ctg cca cag cgc gct gag aac aag ctg cac gtg tcc ctc atg gag     1595
aac tac cca ggc acc ctg gag gcc ctg ggc gag ccc atc cgc cag ggc     1643
cac ttc atc gtg tgg gag ggt gca ccg ggg gcc cgc atg ccc tgg aag     1691
ggc cac aac cgt cac gtg ttc ctc ttc cgc aac cac ctg gta atc tgc     1739
```

```
aag ccc cgg cga gac tcc cgc acc gat acc gtc agc tac gtg ttc cgg     1787
aac atg atg aag ctg agc agc atc gac ctg aac gac cag gtg gag ggg     1835
gat gac cgc gcc ttc gag gtg tgg cag gag cgg gag gac tcg gtg cgc     1883
aag tac ctg ctg cag gca cgg aca gcc att atc aag agc tcg tgg gtg     1931
aag gag atc tgt ggc atc cag cag cgt ctg gcc ctg cct gtg tgg cgg     1979
ccc ccg gac ttt gaa gag gag ctg gcc gac tgc aca gcc gag ctg ggt     2027
gag aca gtc aag ctg gcc tgc cgc gtg acg ggc aca ccc aag cct gtc     2075
atc agc tgg tac aaa gat ggg aaa gca gtg cag gtg gac ccc cac cac     2123
atc ctc att gaa gac cct gat ggc tcg tgt gca ctc atc ctg gac agc     2171
ctg acc ggt gtg gac tct ggc cag tac atg tgc ttc gcg gcc agc gcc     2219
gct ggc aac tgc agt acc ctg ggc aag atc ctg gtg caa gtc cca cca     2267
cgg ttc gtg aac aag gtc cgg gcc tca ccc ttt gtg gag gga gag gac     2315
gcc cag ttc acc tgc acc atc gaa ggc gcc ccg tac ccg cag atc agg     2363
tgg tac aag gac ggg gcc ctg ctg acc act ggc aac aag ttc cag aca     2411
ctg agt gag cct cgc agc ggc ctg cta gtg ctg gtg atc cgg gcg gcc     2459
agc aag gag gac ctg ggg ctc tac gag tgt gag ctg gtg aac cgg ctg     2507
ggc tcc gcg cgg gct agt gcg gag ctg cgc att cag agc ccc atg ctg     2555
cag gcc cag gag cag tgt cac agg gag cag ctc gtg gct gca gtg gaa     2603
gac acc acc ctg gag cga gcg gac cag gag gtc aca tct gtc ctg aag     2651
aga ctg ctg ggc ccc aag gcg cca ggc ccc tcc aca ggg gac ctc act     2699
ggc cct ggc ccc tgc ccc agg ggg gca ccc gca ctc cag gaa acc ggc     2747
tcc cag ccc cca gtc acc gga act tcg gag gca cct gcc gtg ccc ccg     2795
agg gtg cca cag ccc ctc ctc cac gaa ggc cca gag cag gag ccg gag     2843
gcc att gcc aga gcc cag gaa tgg act gtg ccc att cgg atg gag ggt     2891
gca gcc tgg ccc ggg gca ggc aca ggg gag ctg ctc tgg gac gtc cac     2939
agc cac gtg gtc aga gag acc aca cag agg acc tac aca tac cag gcc     2987
atc gac acg cac acc gca cgg ccc cca tcc atg cag gta acc atc gag     3035
gat gtg cag gca cag aca ggc gga acg gcc caa ttc gag gct atc att     3083
gag ggc gac cca cag ccc tcg gtg acc tgg tac aag gac agc gtc cag     3131
ctg gtg gac agc acc cgg ctt agc cag cag caa gaa ggc acc aca tac     3179
tcc ctg gtg ctg agg cat gtg gcc tcg aag gat gcc ggc gtt tac acc     3227
tgc ctg gcc caa aac act ggt ggc cag gtg ctc tgc aag gca gag ctg     3275
ctg gtg ctt ggg ggg gac aat gag ccg gac tca gag aag caa agc cac     3323
cgg agg aag ctg cac tcc ttc tat gag gtc aag gag gag att gga agg     3371
ggc gtg ttt ggc ttc gta aaa aga gtg cag cac aaa gga aac aag atc     3419
ttg tgc gct gcc aag ttc atc ccc cta cgg agc aga act cgg gcc cag     3467
gca tac agg gag cga gac atc ctg gcc gcg ctg agc cac ccg ctg gtc     3515
acg ggg ctg ctg gac cag ttt gag acc cgc aag acc ctc atc ctc atc     3563
ctg gag ctg tgc tca tcc gag gag ctg ctg gac cgc ctg tac agg aag     3611
ggc gtg gtg acg gag gcc gag gtc aag gtc tac atc cag cag ctg gtg     3659
gag ggg ctg cac tac ctg cac agc cat ggc gtt ctc cac ctg gac ata     3707
aag ccc tct aac atc ctg atg gtg cat cct gcc cgg gaa gac att aaa     3755
atc tgc gac ttt ggc ttt gcc cag aac atc acc cca gca gag ctg cag     3803
ttc agc cag tac ggc tcc cct gag ttc gtc tcc ccc gag atc atc cag     3851
cag aac cct gtg agc gaa gcc tcc gac att tgg gcc atg ggt gtc atc     3899
tcc tac ctc agc ctg acc tgc tca tcc cca ttt gcc ggc gag agt gac     3947
cgt gcc acc ctc ctg aac gtc ctg gag ggg cgc gtg tca tgg agc agc     3995
ccc atg gct gcc cac ctc agc gaa gac gcc aaa gac ttc atc aag gct     4043
acg ctg cag aga gcc cct cag gcc cgg cct agt gcg gcc cag tgc ctc     4091
tcc cac ccc tgg ttc ctg aaa tcc atg cct gcg gag gag gcc cac ttc     4139
atc aac acc aag cag ctc aag ttc ctc ctg gcc cga agt cgc tgg cag     4187
cgt tcc ctg atg agc tac aag tcc atc ctg gtg atg cgc tcc atc cct     4235
gag ctg ctg cgg ggc cca ccc gac agc ccc tcc ctc ggc gta gcc cgg     4283
cac ctc tgc agg gac act ggt ggc tcc tcc agt tcc tcc tcc tcc tct     4331
gac aac gag ctc gcc cca ttt gcc cgg gct aag tca ctg cca ccc tcc     4379
ccg gtg aca cac tca cca ctg ctg cac ccc cgg ggc ttc ctg cgg ccc     4427
tcg gcc agc ctg cct gag gaa gcc gag gcc agt gag cgc tcc acc gag     4475
```

```
gcc cca gct ccg cct gca tct ccc gag ggt gcc ggg cca ccg gcc gcc     4523
cag ggc tgc gtg ccc cgg cac agc gtc atc cgc agc ctg ttc tac cac     4571
cag gcg ggt gag agc cct gag cac ggg gcc ctg gcc ccg ggg agc agg     4619
cgg cac ccg gcc cgg cgg cgg cac ctg ctg aag ggc ggg tac att gcg     4667
ggg gcg ctg cca ggc ctg cgc gag cca ctg atg gag cac cgc gtg ctg     4715
gag gag gag gcc gcc agg gag gag cag gcc acc ctc ctg gcc aaa gcc     4763
ccc tca ttc gag act gcc ctc cgg ctg cct gcc tct ggc acc cac ttg     4811
gcc cct ggc cac agc cac tcc ctg gaa cat gac tct ccg agc acc ccc     4859
cgc ccc tcc tcg gag gcc tgc ggt gag gca cag cga ctg cct tca gcc     4907
ccc tcc ggg ggg gcc cct atc agg gac atg ggg cac cct cag ggc tcc     4955
aag cag ctt cca tcc act ggt ggc cac cca ggc act gct cag cca gag     5003
agg cca tcc ccg gac agc cct tgg ggg cag cca gcc cct ttc tgc cac     5051
ccc aag cag ggt tct gcc ccc cag gag ggc tgc agc ccc cac cca gca     5099
gtt gcc cca tgc cct cct ggc tcc ttc cct cca gga tct tgc aaa gag     5147
gcc ccc tta gta ccc tca agc ccc ttc ttg gga cag ccc cag gca ccc     5195
cct gcc cct gcc aaa gca agc ccc cca ttg gac tct aag atg ggg cct     5243
gga gac atc tct ctt cct ggg agg cca aaa ccc ggc ccc tgc agt tcc     5291
cca ggg tca gcc tcc cag gcg agc tct tcc caa gtg agc tcc ctc agg     5339
gtg ggc tcc tcc cag gtg ggc aca gag cct ggc ccc tcc ctg gat gcg     5387
gag ggc tgg acc cag gag gct gag gat ctg tcc gac tcc aca ccc acc     5435
ttg cag cgg cct cag gaa cag gcg acc atg cgc aag ttc tcc ctg ggt     5483
ggt cgc ggg ggc tac gca ggc gtg gct ggc tat ggc acc ttt gcc ttt     5531
ggt gga gat gca ggg ggc atg ctg ggg cag ggg ccc atg tgg gcc agg     5579
ata gcc tgg gct gtg tcc cag tca gag gag gag gag cag gag gag gcc     5627
agg gct gag tcc cag tcg gag gag cag cag gag gcc agg gct gag agc     5675
cca ctg ccc cag gtc agt gca agg cct gtg cct gag gtc ggc agg gct     5723
ccc acc agg agc tct cca gag ccc acc cca tgg gag gac atc ggg cag     5771
gtc tcc ctg gtg cag atc cgg gac ctg tca ggt gat gcg gag gcg gcc     5819
gac aca ata tcc ctg gac att tcc gag gtg gac ccc gcc tac ctc aac     5867
ctc tca gac ctg tac gat atc aag tac ctc cca ttc gag ttt atg atc     5915
ttc agg aaa gtc ccc aag tcc gct cag cca gag ccg ccc tcc ccc atg     5963
gct gag gag gag ctg gcc gag ttc ccg gag ccc acg tgg ccc tgg cca     6011
ggt gaa ctg ggc ccc cac gca ggc ctg gag atc aca gag gag tca gag     6059
gat gtg gac gcg ctg ctg gca gag gct gcc gtg ggc agg aag cgc aag     6107
tgg tcc tcg ccg tca cgc agc ctc ttc cac ttc cct ggg agg cac ctg     6155
ccg ctg gat gag cct gca gag ctg ggg ctg cgt gag aga gtg aag gcc     6203
tcc gtg gag cac atc tcc cgg atc ctg aag ggc agg ccg gaa ggt ctg     6251
gag aag gag ggg ccc ccc agg aag aag cca ggc ctt gct tcc ttc cgg     6299
ctc tca ggt ctg aag agc tgg gac cga gcg ccg aca ttc cta agg gag     6347
ctc tca gat gag act gtg gtc ctg ggc cag tca gtg aca ctg gcc tgc     6395
cag gtg tca gcc cag cca gct gcc cag gcc acc tgg agc aaa gac gga     6443
gcc ccc ctg gag agc agc agc cgt gtc ctc atc tct gcc acc ctc aag     6491
aac ttc cag ctt ctg acc atc ctg gtg gtg gtg gct gag gac ctg ggt     6539
gtg tac acc tgc agc gtg agc aat gcg ctg ggg aca gtg acc acc acg     6587
ggc gtc ctc cgg aag gca gag cgc ccc tca tct tcg cca tgc ccg gat     6635
atc ggg gag gtg tac gcg gat ggg gtg ctg ctg gtc tgg aag ccc gtg     6683
gaa tcc tac ggc cct gtg acc tac att gtg cag tgc agc cta gaa ggc     6731
ggc agc tgg acc aca ctg gcc tcc gac atc ttt gac tgc tgc tac ctg     6779
acc agc aag ctc tcc cgg ggt ggc acc tac acc ttc cgc acg gca tgt     6827
gtc agc aag gca gga atg ggt ccc tac agc agc ccc tcg gag caa gtc     6875
ctc ctg gga gcg ccc agc cac ctg gcc tct gag gag gag agc cag ggg     6923
cgg tca gcc caa ccc ctg ccc agc aca aag acc ttc gca ttc cag aca     6971
cag atc cag agg ggc cgc ttc agc gtg gtg cgg caa tgc tgg gag aag     7019
gcc agc ggg cgg gcg ctg gcc gcc aag atc atc ccc tac cac ccc aag     7067
gac aag aca gca gtg ctg cgc gaa tac gag gcc ctc aag ggc ctg cgc     7115
cac ccg cac ctg gcc cag ctg cac gca gcc tac ctc agc ccc cgg cac     7163
ctg gtg ctc atc ttg gag ctg tgc tct ggg ccc gag ctg ctc ccc tgc     7211
```

```
ctg gcc gag agg gcc tcc tac tca gaa tcc gag gtg aag gac tac ctg     7259
tgg cag atg ttg agt gcc acc cag tac ctg cac aac cag cac atc ctg     7307
cac ctg gac ctg agg tcc gag aac atg atc atc acc gaa tac aac ctg     7355
ctc aag gtc gtg gac ctg ggc aat gca cag agc ctc agc cag gag aag     7403
gtg ctg ccc tca gac aag ttc aag gac tac cta gag acc atg gct cca     7451
gag ctc ctg gag ggc cag ggg gct gtt cca cag aca gac atc tgg gcc     7499
atc ggt gtg aca gcc ttc atc atg ctg agc gcc gag tac ccg gtg agc     7547
agc gag ggt gca cgc gac ctg cag aga gga ctg cgc aag ggg ctg gtc     7595
cgg ctg agc cgc tgc tac gcg ggg ctg tcc ggg ggc gcc gtg gcc ttc     7643
ctg cgc agc act ctg tgc gcc cag ccc tgg ggc cgg ccc tgc gcg tcc     7691
agc tgc ctg cag tgc ccg tgg cta aca gag gag ggc ccg gcc tgt tcg     7739
cgg ccc gcg ccc gtg acc ttc cct acc gcg cgg ctg cgc gtc ttc gtg     7787
cgc aat cgc gag aag aga cgc gcg ctg ctg tac aag agg cac aac ctg     7835
gcc cag gtg cgc tgagggtcgcc ccggccacac ccttggtctc cccgctgggg       7888
gtcgctgcag acgcgccaat aaaaacgcac agccgggcga                        7928
```

<210> 6
<211> 2596
<212> PRT
<213> Homo sapiens

<400> 6

```
Met Leu Glu Arg Phe Thr Pro Lys Lys Val Lys Lys Gly Ser Ser Ile
 1               5                  10                  15
Thr Phe Ser Val Lys Val Glu Gly Arg Pro Val Pro Thr Val His Trp
            20                  25                  30
Leu Arg Glu Glu Ala Glu Arg Gly Val Leu Trp Ile Gly Pro Asp Thr
        35                  40                  45
Pro Gly Tyr Thr Val Ala Ser Ser Ala Gln Gln His Ser Leu Val Leu
    50                  55                  60
Leu Asp Val Gly Arg Gln His Gln Gly Thr Tyr Thr Cys Ile Ala Ser
65                  70                  75                  80
Asn Ala Ala Gly Gln Ala Leu Cys Ser Ala Ser Leu His Val Ser Gly
                85                  90                  95
Leu Pro Lys Val Glu Glu Gln Glu Lys Val Lys Glu Ala Leu Ile Ser
            100                 105                 110
Thr Phe Leu Gln Gly Thr Thr Gln Ala Ile Ser Ala Gln Gly Leu Glu
        115                 120                 125
Thr Ala Ser Phe Ala Asp Leu Gly Gly Gln Arg Lys Glu Glu Pro Leu
    130                 135                 140
Ala Ala Lys Glu Ala Leu Gly His Leu Ser Leu Ala Glu Val Gly Thr
145                 150                 155                 160
Glu Glu Phe Leu Gln Lys Leu Thr Ser Gln Ile Thr Glu Met Val Ser
                165                 170                 175
Ala Lys Ile Thr Gln Ala Lys Leu Gln Val Pro Gly Gly Asp Ser Asp
            180                 185                 190
Glu Asp Ser Lys Thr Pro Ser Ala Ser Pro Arg His Gly Arg Ser Arg
        195                 200                 205
Pro Ser Ser Ser Ile Gln Glu Ser Ser Ser Glu Ser Glu Asp Gly Asp
    210                 215                 220
Ala Arg Gly Glu Ile Phe Asp Ile Tyr Val Val Thr Ala Asp Tyr Leu
225                 230                 235                 240
```

```
Pro Leu Gly Ala Glu Gln Asp Ala Ile Thr Leu Arg Glu Gly Gln Tyr
                245                 250                 255
Val Glu Val Leu Asp Ala Ala His Pro Leu Arg Trp Leu Val Arg Thr
            260                 265                 270
Lys Pro Thr Lys Ser Ser Pro Ser Arg Gln Gly Trp Val Ser Pro Ala
        275                 280                 285
Tyr Leu Asp Arg Arg Leu Lys Leu Ser Pro Glu Trp Gly Ala Ala Glu
    290                 295                 300
Ala Pro Glu Phe Pro Gly Glu Ala Val Ser Glu Asp Glu Tyr Lys Ala
305                 310                 315                 320
Arg Leu Ser Ser Val Ile Gln Glu Leu Leu Ser Ser Glu Gln Ala Phe
                325                 330                 335
Val Glu Glu Leu Gln Phe Leu Gln Ser His His Leu Gln His Leu Glu
            340                 345                 350
Arg Cys Pro His Val Pro Ile Ala Val Ala Gly Gln Lys Ala Val Ile
        355                 360                 365
Phe Arg Asn Val Arg Asp Ile Gly Arg Phe His Ser Ser Phe Leu Gln
    370                 375                 380
Glu Leu Gln Gln Cys Asp Thr Asp Asp Asp Val Ala Met Cys Phe Ile
385                 390                 395                 400
Lys Asn Gln Ala Ala Phe Glu Gln Tyr Leu Glu Phe Leu Val Gly Arg
                405                 410                 415
Val Gln Ala Glu Ser Val Val Val Ser Thr Ala Ile Gln Glu Phe Tyr
            420                 425                 430
Lys Lys Tyr Ala Glu Glu Ala Leu Leu Ala Gly Asp Pro Ser Gln Pro
        435                 440                 445
Pro Pro Pro Pro Leu Gln His Tyr Leu Glu Gln Pro Val Glu Arg Val
    450                 455                 460
Gln Arg Tyr Gln Ala Leu Leu Lys Glu Leu Ile Arg Asn Lys Ala Arg
465                 470                 475                 480
Asn Arg Gln Asn Cys Ala Leu Leu Glu Gln Ala Tyr Ala Val Val Ser
                485                 490                 495
Ala Leu Pro Gln Arg Ala Glu Asn Lys Leu His Val Ser Leu Met Glu
            500                 505                 510
Asn Tyr Pro Gly Thr Leu Glu Ala Leu Gly Glu Pro Ile Arg Gln Gly
        515                 520                 525
His Phe Ile Val Trp Glu Gly Ala Pro Gly Ala Arg Met Pro Trp Lys
    530                 535                 540
Gly His Asn Arg His Val Phe Leu Phe Arg Asn His Leu Val Ile Cys
545                 550                 555                 560
Lys Pro Arg Arg Asp Ser Arg Thr Asp Thr Val Ser Tyr Val Phe Arg
                565                 570                 575
Asn Met Met Lys Leu Ser Ser Ile Asp Leu Asn Asp Gln Val Glu Gly
            580                 585                 590
Asp Asp Arg Ala Phe Glu Val Trp Gln Glu Arg Glu Asp Ser Val Arg
        595                 600                 605
Lys Tyr Leu Leu Gln Ala Arg Thr Ala Ile Ile Lys Ser Ser Trp Val
    610                 615                 620
Lys Glu Ile Cys Gly Ile Gln Gln Arg Leu Ala Leu Pro Val Trp Arg
625                 630                 635                 640
Pro Pro Asp Phe Glu Glu Glu Leu Ala Asp Cys Thr Ala Glu Leu Gly
                645                 650                 655
Glu Thr Val Lys Leu Ala Cys Arg Val Thr Gly Thr Pro Lys Pro Val
            660                 665                 670
Ile Ser Trp Tyr Lys Asp Gly Lys Ala Val Gln Val Asp Pro His His
        675                 680                 685
Ile Leu Ile Glu Asp Pro Asp Gly Ser Cys Ala Leu Ile Leu Asp Ser
```

```
    690                 695                 700
Leu Thr Gly Val Asp Ser Gly Gln Tyr Met Cys Phe Ala Ala Ser Ala
705                 710                 715                 720
Ala Gly Asn Cys Ser Thr Leu Gly Lys Ile Leu Val Gln Val Pro Pro
                725                 730                 735
Arg Phe Val Asn Lys Val Arg Ala Ser Pro Phe Val Glu Gly Glu Asp
            740                 745                 750
Ala Gln Phe Thr Cys Thr Ile Glu Gly Ala Pro Tyr Pro Gln Ile Arg
        755                 760                 765
Trp Tyr Lys Asp Gly Ala Leu Leu Thr Thr Gly Asn Lys Phe Gln Thr
    770                 775                 780
Leu Ser Glu Pro Arg Ser Gly Leu Leu Val Leu Val Ile Arg Ala Ala
785                 790                 795                 800
Ser Lys Glu Asp Leu Gly Leu Tyr Glu Cys Glu Leu Val Asn Arg Leu
                805                 810                 815
Gly Ser Ala Arg Ala Ser Ala Glu Leu Arg Ile Gln Ser Pro Met Leu
            820                 825                 830
Gln Ala Gln Glu Gln Cys His Arg Glu Gln Leu Val Ala Ala Val Glu
        835                 840                 845
Asp Thr Thr Leu Glu Arg Ala Asp Gln Glu Val Thr Ser Val Leu Lys
    850                 855                 860
Arg Leu Leu Gly Pro Lys Ala Pro Gly Pro Ser Thr Gly Asp Leu Thr
865                 870                 875                 880
Gly Pro Gly Pro Cys Pro Arg Gly Ala Pro Ala Leu Gln Glu Thr Gly
                885                 890                 895
Ser Gln Pro Pro Val Thr Gly Thr Ser Glu Ala Pro Ala Val Pro Pro
            900                 905                 910
Arg Val Pro Gln Pro Leu Leu His Glu Gly Pro Glu Gln Glu Pro Glu
        915                 920                 925
Ala Ile Ala Arg Ala Gln Glu Trp Thr Val Pro Ile Arg Met Glu Gly
    930                 935                 940
Ala Ala Trp Pro Gly Ala Gly Thr Gly Glu Leu Leu Trp Asp Val His
945                 950                 955                 960
Ser His Val Val Arg Glu Thr Thr Gln Arg Thr Tyr Thr Tyr Gln Ala
                965                 970                 975
Ile Asp Thr His Thr Ala Arg Pro Pro Ser Met Gln Val Thr Ile Glu
            980                 985                 990
Asp Val Gln Ala Gln Thr Gly Gly Thr Ala Gln Phe Glu Ala Ile Ile
        995                 1000                1005
Glu Gly Asp Pro Gln Pro Ser Val Thr Trp Tyr Lys Asp Ser Val Gln
    1010                1015                1020
Leu Val Asp Ser Thr Arg Leu Ser Gln Gln Gln Glu Gly Thr Thr Tyr
1025                1030                1035                1040
Ser Leu Val Leu Arg His Val Ala Ser Lys Asp Ala Gly Val Tyr Thr
                1045                1050                1055
Cys Leu Ala Gln Asn Thr Gly Gly Gln Val Leu Cys Lys Ala Glu Leu
            1060                1065                1070
Leu Val Leu Gly Gly Asp Asn Glu Pro Asp Ser Glu Lys Gln Ser His
        1075                1080                1085
Arg Arg Lys Leu His Ser Phe Tyr Glu Val Lys Glu Glu Ile Gly Arg
    1090                1095                1100
Gly Val Phe Gly Phe Val Lys Arg Val Gln His Lys Gly Asn Lys Ile
1105                1110                1115                1120
Leu Cys Ala Ala Lys Phe Ile Pro Leu Arg Ser Arg Thr Arg Ala Gln
                1125                1130                1135
Ala Tyr Arg Glu Arg Asp Ile Leu Ala Ala Leu Ser His Pro Leu Val
            1140                1145                1150
```

```
Thr Gly Leu Leu Asp Gln Phe Glu Thr Arg Lys Thr Leu Ile Leu Ile
        1155                1160                1165
Leu Glu Leu Cys Ser Ser Glu Glu Leu Leu Asp Arg Leu Tyr Arg Lys
    1170                1175                1180
Gly Val Val Thr Glu Ala Glu Val Lys Val Tyr Ile Gln Gln Leu Val
1185                1190                1195                1200
Glu Gly Leu His Tyr Leu His Ser His Gly Val Leu His Leu Asp Ile
                1205                1210                1215
Lys Pro Ser Asn Ile Leu Met Val His Pro Ala Arg Glu Asp Ile Lys
            1220                1225                1230
Ile Cys Asp Phe Gly Phe Ala Gln Asn Ile Thr Pro Ala Glu Leu Gln
        1235                1240                1245
Phe Ser Gln Tyr Gly Ser Pro Glu Phe Val Ser Pro Glu Ile Ile Gln
    1250                1255                1260
Gln Asn Pro Val Ser Glu Ala Ser Asp Ile Trp Ala Met Gly Val Ile
1265                1270                1275                1280
Ser Tyr Leu Ser Leu Thr Cys Ser Ser Pro Phe Ala Gly Glu Ser Asp
                1285                1290                1295
Arg Ala Thr Leu Leu Asn Val Leu Glu Gly Arg Val Ser Trp Ser Ser
            1300                1305                1310
Pro Met Ala Ala His Leu Ser Glu Asp Ala Lys Asp Phe Ile Lys Ala
        1315                1320                1325
Thr Leu Gln Arg Ala Pro Gln Ala Arg Pro Ser Ala Ala Gln Cys Leu
    1330                1335                1340
Ser His Pro Trp Phe Leu Lys Ser Met Pro Ala Glu Glu Ala His Phe
1345                1350                1355                1360
Ile Asn Thr Lys Gln Leu Lys Phe Leu Leu Ala Arg Ser Arg Trp Gln
                1365                1370                1375
Arg Ser Leu Met Ser Tyr Lys Ser Ile Leu Val Met Arg Ser Ile Pro
            1380                1385                1390
Glu Leu Leu Arg Gly Pro Pro Asp Ser Pro Ser Leu Gly Val Ala Arg
        1395                1400                1405
His Leu Cys Arg Asp Thr Gly Gly Ser Ser Ser Ser Ser Ser Ser Ser
    1410                1415                1420
Asp Asn Glu Leu Ala Pro Phe Ala Arg Ala Lys Ser Leu Pro Pro Ser
1425                1430                1435                1440
Pro Val Thr His Ser Pro Leu Leu His Pro Arg Gly Phe Leu Arg Pro
                1445                1450                1455
Ser Ala Ser Leu Pro Glu Glu Ala Glu Ala Ser Glu Arg Ser Thr Glu
            1460                1465                1470
Ala Pro Ala Pro Pro Ala Ser Pro Glu Gly Ala Gly Pro Pro Ala Ala
        1475                1480                1485
Gln Gly Cys Val Pro Arg His Ser Val Ile Arg Ser Leu Phe Tyr His
    1490                1495                1500
Gln Ala Gly Glu Ser Pro Glu His Gly Ala Leu Ala Pro Gly Ser Arg
1505                1510                1515                1520
Arg His Pro Ala Arg Arg Arg His Leu Leu Lys Gly Gly Tyr Ile Ala
                1525                1530                1535
Gly Ala Leu Pro Gly Leu Arg Glu Pro Leu Met Glu His Arg Val Leu
            1540                1545                1550
Glu Glu Glu Ala Ala Arg Glu Glu Gln Ala Thr Leu Leu Ala Lys Ala
        1555                1560                1565
Pro Ser Phe Glu Thr Ala Leu Arg Leu Pro Ala Ser Gly Thr His Leu
    1570                1575                1580
Ala Pro Gly His Ser His Ser Leu Glu His Asp Ser Pro Ser Thr Pro
1585                1590                1595                1600
Arg Pro Ser Ser Glu Ala Cys Gly Glu Ala Gln Arg Leu Pro Ser Ala
```

```
                1605                1610                1615
Pro Ser Gly Gly Ala Pro Ile Arg Asp Met Gly His Pro Gln Gly Ser
            1620                1625                1630
Lys Gln Leu Pro Ser Thr Gly Gly His Pro Gly Thr Ala Gln Pro Glu
        1635                1640                1645
Arg Pro Ser Pro Asp Ser Pro Trp Gly Gln Pro Ala Pro Phe Cys His
    1650                1655                1660
Pro Lys Gln Gly Ser Ala Pro Gln Glu Gly Cys Ser Pro His Pro Ala
1665                1670                1675                1680
Val Ala Pro Cys Pro Pro Gly Ser Phe Pro Pro Gly Ser Cys Lys Glu
                1685                1690                1695
Ala Pro Leu Val Pro Ser Ser Pro Phe Leu Gly Gln Pro Gln Ala Pro
            1700                1705                1710
Pro Ala Pro Ala Lys Ala Ser Pro Pro Leu Asp Ser Lys Met Gly Pro
        1715                1720                1725
Gly Asp Ile Ser Leu Pro Gly Arg Pro Lys Pro Gly Pro Cys Ser Ser
    1730                1735                1740
Pro Gly Ser Ala Ser Gln Ala Ser Ser Ser Gln Val Ser Ser Leu Arg
1745                1750                1755                1760
Val Gly Ser Ser Gln Val Gly Thr Glu Pro Gly Pro Ser Leu Asp Ala
                1765                1770                1775
Glu Gly Trp Thr Gln Glu Ala Glu Asp Leu Ser Asp Ser Thr Pro Thr
            1780                1785                1790
Leu Gln Arg Pro Gln Glu Gln Ala Thr Met Arg Lys Phe Ser Leu Gly
        1795                1800                1805
Gly Arg Gly Gly Tyr Ala Gly Val Ala Gly Tyr Gly Thr Phe Ala Phe
    1810                1815                1820
Gly Gly Asp Ala Gly Gly Met Leu Gly Gln Gly Pro Met Trp Ala Arg
1825                1830                1835                1840
Ile Ala Trp Ala Val Ser Gln Ser Glu Glu Glu Glu Gln Glu Glu Ala
                1845                1850                1855
Arg Ala Glu Ser Gln Ser Glu Glu Gln Gln Glu Ala Arg Ala Glu Ser
            1860                1865                1870
Pro Leu Pro Gln Val Ser Ala Arg Pro Val Pro Glu Val Gly Arg Ala
        1875                1880                1885
Pro Thr Arg Ser Ser Pro Glu Pro Thr Pro Trp Glu Asp Ile Gly Gln
    1890                1895                1900
Val Ser Leu Val Gln Ile Arg Asp Leu Ser Gly Asp Ala Glu Ala Ala
1905                1910                1915                1920
Asp Thr Ile Ser Leu Asp Ile Ser Glu Val Asp Pro Ala Tyr Leu Asn
                1925                1930                1935
Leu Ser Asp Leu Tyr Asp Ile Lys Tyr Leu Pro Phe Glu Phe Met Ile
            1940                1945                1950
Phe Arg Lys Val Pro Lys Ser Ala Gln Pro Glu Pro Pro Ser Pro Met
        1955                1960                1965
Ala Glu Glu Glu Leu Ala Glu Phe Pro Glu Pro Thr Trp Pro Trp Pro
    1970                1975                1980
Gly Glu Leu Gly Pro His Ala Gly Leu Glu Ile Thr Glu Glu Ser Glu
1985                1990                1995                2000
Asp Val Asp Ala Leu Leu Ala Glu Ala Ala Val Gly Arg Lys Arg Lys
                2005                2010                2015
Trp Ser Ser Pro Ser Arg Ser Leu Phe His Phe Pro Gly Arg His Leu
            2020                2025                2030
Pro Leu Asp Glu Pro Ala Glu Leu Gly Leu Arg Glu Arg Val Lys Ala
        2035                2040                2045
Ser Val Glu His Ile Ser Arg Ile Leu Lys Gly Arg Pro Glu Gly Leu
    2050                2055                2060
```

```
Glu Lys Glu Gly Pro Pro Arg Lys Lys Pro Gly Leu Ala Ser Phe Arg
2065                2070                2075                2080
Leu Ser Gly Leu Lys Ser Trp Asp Arg Ala Pro Thr Phe Leu Arg Glu
                    2085                2090                2095
Leu Ser Asp Glu Thr Val Val Leu Gly Gln Ser Val Thr Leu Ala Cys
            2100                2105                2110
Gln Val Ser Ala Gln Pro Ala Ala Gln Ala Thr Trp Ser Lys Asp Gly
        2115                2120                2125
Ala Pro Leu Glu Ser Ser Ser Arg Val Leu Ile Ser Ala Thr Leu Lys
    2130                2135                2140
Asn Phe Gln Leu Leu Thr Ile Leu Val Val Val Ala Glu Asp Leu Gly
2145                2150                2155                2160
Val Tyr Thr Cys Ser Val Ser Asn Ala Leu Gly Thr Val Thr Thr Thr
                    2165                2170                2175
Gly Val Leu Arg Lys Ala Glu Arg Pro Ser Ser Ser Pro Cys Pro Asp
            2180                2185                2190
Ile Gly Glu Val Tyr Ala Asp Gly Val Leu Leu Val Trp Lys Pro Val
        2195                2200                2205
Glu Ser Tyr Gly Pro Val Thr Tyr Ile Val Gln Cys Ser Leu Glu Gly
    2210                2215                2220
Gly Ser Trp Thr Thr Leu Ala Ser Asp Ile Phe Asp Cys Cys Tyr Leu
2225                2230                2235                2240
Thr Ser Lys Leu Ser Arg Gly Gly Thr Tyr Thr Phe Arg Thr Ala Cys
                    2245                2250                2255
Val Ser Lys Ala Gly Met Gly Pro Tyr Ser Ser Pro Ser Glu Gln Val
            2260                2265                2270
Leu Leu Gly Ala Pro Ser His Leu Ala Ser Glu Glu Glu Ser Gln Gly
        2275                2280                2285
Arg Ser Ala Gln Pro Leu Pro Ser Thr Lys Thr Phe Ala Phe Gln Thr
    2290                2295                2300
Gln Ile Gln Arg Gly Arg Phe Ser Val Val Arg Gln Cys Trp Glu Lys
2305                2310                2315                2320
Ala Ser Gly Arg Ala Leu Ala Ala Lys Ile Ile Pro Tyr His Pro Lys
                    2325                2330                2335
Asp Lys Thr Ala Val Leu Arg Glu Tyr Glu Ala Leu Lys Gly Leu Arg
            2340                2345                2350
His Pro His Leu Ala Gln Leu His Ala Ala Tyr Leu Ser Pro Arg His
        2355                2360                2365
Leu Val Leu Ile Leu Glu Leu Cys Ser Gly Pro Glu Leu Leu Pro Cys
    2370                2375                2380
Leu Ala Glu Arg Ala Ser Tyr Ser Glu Ser Glu Val Lys Asp Tyr Leu
2385                2390                2395                2400
Trp Gln Met Leu Ser Ala Thr Gln Tyr Leu His Asn Gln His Ile Leu
                    2405                2410                2415
His Leu Asp Leu Arg Ser Glu Asn Met Ile Ile Thr Glu Tyr Asn Leu
            2420                2425                2430
Leu Lys Val Val Asp Leu Gly Asn Ala Gln Ser Leu Ser Gln Glu Lys
        2435                2440                2445
Val Leu Pro Ser Asp Lys Phe Lys Asp Tyr Leu Glu Thr Met Ala Pro
    2450                2455                2460
Glu Leu Leu Glu Gly Gln Gly Ala Val Pro Gln Thr Asp Ile Trp Ala
2465                2470                2475                2480
Ile Gly Val Thr Ala Phe Ile Met Leu Ser Ala Glu Tyr Pro Val Ser
                    2485                2490                2495
Ser Glu Gly Ala Arg Asp Leu Gln Arg Gly Leu Arg Lys Gly Leu Val
            2500                2505                2510
Arg Leu Ser Arg Cys Tyr Ala Gly Leu Ser Gly Gly Ala Val Ala Phe
```

```
        2515                2520                2525
Leu Arg Ser Thr Leu Cys Ala Gln Pro Trp Gly Arg Pro Cys Ala Ser
    2530                2535                2540
Ser Cys Leu Gln Cys Pro Trp Leu Thr Glu Glu Gly Pro Ala Cys Ser
2545                2550                2555                2560
Arg Pro Ala Pro Val Thr Phe Pro Thr Ala Arg Leu Arg Val Phe Val
                2565                2570                2575
Arg Asn Arg Glu Lys Arg Arg Ala Leu Leu Tyr Lys Arg His Asn Leu
            2580                2585                2590
Ala Gln Val Arg
        2595
```

<210> 7
<211> 871
<212> PRT
<213> Homo sapiens

<400> 7

```
Met Gly Pro Gly Asp Ile Ser Leu Pro Gly Arg Pro Lys Pro Gly Pro
 1               5                  10                  15
Cys Ser Ser Pro Gly Ser Ala Ser Gln Ala Ser Ser Ser Gln Val Ser
            20                  25                  30
Ser Leu Arg Val Gly Ser Ser Gln Val Gly Thr Glu Pro Gly Pro Ser
        35                  40                  45
Leu Asp Ala Glu Gly Trp Thr Gln Glu Ala Glu Asp Leu Ser Asp Ser
    50                  55                  60
Thr Pro Thr Leu Gln Arg Pro Gln Glu Gln Ala Thr Met Arg Lys Phe
65                  70                  75                  80
Ser Leu Gly Gly Arg Gly Gly Tyr Ala Gly Val Ala Gly Tyr Gly Thr
                85                  90                  95
Phe Ala Phe Gly Gly Asp Ala Gly Gly Met Leu Gly Gln Gly Pro Met
            100                 105                 110
Trp Ala Arg Ile Ala Trp Ala Val Ser Gln Ser Glu Glu Glu Glu Gln
        115                 120                 125
Glu Glu Ala Arg Ala Glu Ser Gln Ser Glu Glu Gln Gln Glu Ala Arg
    130                 135                 140
Ala Glu Ser Pro Leu Pro Gln Val Ser Ala Arg Pro Val Pro Glu Val
145                 150                 155                 160
Gly Arg Ala Pro Thr Arg Ser Ser Pro Glu Pro Thr Pro Trp Glu Asp
                165                 170                 175
Ile Gly Gln Val Ser Leu Val Gln Ile Arg Asp Leu Ser Gly Asp Ala
            180                 185                 190
Glu Ala Ala Asp Thr Ile Ser Leu Asp Ile Ser Glu Val Asp Pro Ala
        195                 200                 205
Tyr Leu Asn Leu Ser Asp Leu Tyr Asp Ile Lys Tyr Leu Pro Phe Glu
    210                 215                 220
Phe Met Ile Phe Arg Lys Val Pro Lys Ser Ala Gln Pro Glu Pro Pro
225                 230                 235                 240
Ser Pro Met Ala Glu Glu Glu Leu Ala Glu Phe Pro Glu Pro Thr Trp
```

```
                245                 250                 255
Pro Trp Pro Gly Glu Leu Gly Pro His Ala Gly Leu Glu Ile Thr Glu
            260                 265                 270
Glu Ser Glu Asp Val Asp Ala Leu Leu Ala Glu Ala Ala Val Gly Arg
        275                 280                 285
Lys Arg Lys Trp Ser Ser Pro Ser Arg Ser Leu Phe His Phe Pro Gly
    290                 295                 300
Arg His Leu Pro Leu Asp Glu Pro Ala Glu Leu Gly Leu Arg Glu Arg
305                 310                 315                 320
Val Lys Ala Ser Val Glu His Ile Ser Arg Ile Leu Lys Gly Arg Pro
                325                 330                 335
Glu Gly Leu Glu Lys Glu Gly Pro Pro Arg Lys Lys Pro Gly Leu Ala
            340                 345                 350
Ser Phe Arg Leu Ser Gly Leu Lys Ser Trp Asp Arg Ala Pro Thr Phe
        355                 360                 365
Leu Arg Glu Leu Ser Asp Glu Thr Val Val Leu Gly Gln Ser Val Thr
    370                 375                 380
Leu Ala Cys Gln Val Ser Ala Gln Pro Ala Ala Gln Ala Thr Trp Ser
385                 390                 395                 400
Lys Asp Gly Ala Pro Leu Glu Ser Ser Ser Arg Val Leu Ile Ser Ala
                405                 410                 415
Thr Leu Lys Asn Phe Gln Leu Leu Thr Ile Leu Val Val Val Ala Glu
            420                 425                 430
Asp Leu Gly Val Tyr Thr Cys Ser Val Ser Asn Ala Leu Gly Thr Val
        435                 440                 445
Thr Thr Thr Gly Val Leu Arg Lys Ala Glu Arg Pro Ser Ser Ser Pro
    450                 455                 460
Cys Pro Asp Ile Gly Glu Val Tyr Ala Asp Gly Val Leu Leu Val Trp
465                 470                 475                 480
Lys Pro Val Glu Ser Tyr Gly Pro Val Thr Tyr Ile Val Gln Cys Ser
                485                 490                 495
Leu Glu Gly Gly Ser Trp Thr Thr Leu Ala Ser Asp Ile Phe Asp Cys
            500                 505                 510
Cys Tyr Leu Thr Ser Lys Leu Ser Arg Gly Gly Thr Tyr Thr Phe Arg
        515                 520                 525
Thr Ala Cys Val Ser Lys Ala Gly Met Gly Pro Tyr Ser Ser Pro Ser
    530                 535                 540
Glu Gln Val Leu Leu Gly Gly Pro Ser His Leu Ala Ser Glu Glu Glu
545                 550                 555                 560
Ser Gln Gly Arg Ser Ala Gln Pro Leu Pro Ser Thr Lys Thr Phe Ala
                565                 570                 575
Phe Gln Thr Gln Ile Gln Arg Gly Arg Phe Ser Val Val Arg Gln Cys
            580                 585                 590
Trp Glu Lys Ala Ser Gly Arg Ala Leu Ala Ala Lys Ile Ile Pro Tyr
        595                 600                 605
His Pro Lys Asp Lys Thr Ala Val Leu Arg Glu Tyr Glu Ala Leu Lys
    610                 615                 620
Gly Leu Arg His Pro His Leu Ala Gln Leu His Ala Ala Tyr Leu Ser
625                 630                 635                 640
Pro Arg His Leu Val Leu Ile Leu Glu Leu Cys Ser Gly Pro Glu Leu
                645                 650                 655
Leu Pro Cys Leu Ala Glu Arg Ala Ser Tyr Ser Glu Ser Glu Val Lys
            660                 665                 670
Asp Tyr Leu Trp Gln Met Leu Ser Ala Thr Gln Tyr Leu His Asn Gln
        675                 680                 685
His Ile Leu His Leu Asp Leu Arg Ser Glu Asn Met Ile Ile Thr Glu
    690                 695                 700
```

```
Tyr Asn Leu Leu Lys Val Val Asp Leu Gly Asn Ala Gln Ser Leu Ser
705                 710                 715                 720
Gln Glu Lys Val Leu Pro Ser Asp Lys Phe Lys Asp Tyr Leu Glu Thr
                725                 730                 735
Met Ala Pro Glu Leu Leu Glu Gly Gln Gly Ala Val Pro Gln Thr Asp
            740                 745                 750
Ile Trp Ala Ile Gly Val Thr Ala Phe Ile Met Leu Ser Ala Glu Tyr
        755                 760                 765
Pro Val Ser Ser Glu Gly Ala Arg Asp Leu Gln Arg Gly Leu Arg Lys
    770                 775                 780
Gly Leu Val Arg Leu Ser Arg Cys Tyr Ala Gly Leu Ser Gly Gly Ala
785                 790                 795                 800
Val Ala Phe Leu Arg Ser Thr Leu Cys Ala Gln Pro Trp Gly Arg Pro
                805                 810                 815
Cys Ala Ser Ser Cys Leu Gln Cys Pro Trp Leu Thr Glu Glu Gly Pro
            820                 825                 830
Ala Cys Ser Arg Pro Ala Pro Val Thr Phe Pro Thr Ala Arg Leu Arg
        835                 840                 845
Val Phe Val Arg Asn Arg Glu Lys Arg Arg Ala Leu Leu Tyr Lys Arg
    850                 855                 860
His Asn Leu Ala Gln Val Arg
865                 870
```

<210> 8
<211> 548
<212> PRT
<213> Rattus norvegicus

<400> 8

```
Met Ala His Ile Ser Arg Ile Leu Lys Gly Lys Pro Glu Gly Pro Glu
 1                5                  10                  15
Lys Glu Gly Pro Pro Arg Lys Lys Ala Gly Leu Ala Ser Phe Arg Leu
            20                  25                  30
Ser Gly Leu Lys Gly Arg Asp Gln Ala Pro Ser Phe Leu Arg Glu Leu
        35                  40                  45
Ser Asp Glu Ala Val Val Leu Gly Gln Ser Val Thr Leu Ala Cys Gln
    50                  55                  60
Val Leu Ala Gln Pro Thr Ala Gln Ala Thr Trp Ser Lys Asp Gly Ala
65                  70                  75                  80
Leu Leu Glu Ser Ser Gly His Leu Leu Ile Ser Ser Thr Leu Lys Asn
                85                  90                  95
Phe Gln Leu Leu Thr Ile Leu Val Val Thr Glu Glu Asp Leu Gly Thr
            100                 105                 110
Tyr Thr Cys Cys Val Ser Asn Pro Leu Gly Thr Ala Val Thr Thr Gly
        115                 120                 125
Val Leu Arg Lys Ala Glu Arg Pro Ser Ser Ser Pro Arg Pro Glu Val
    130                 135                 140
Gly Glu Leu Tyr Thr Asp Ala Val Leu Leu Val Trp Lys Pro Val Glu
145                 150                 155                 160
Ser Tyr Gly Pro Val Thr Tyr Ile Val Gln Cys Cys Ile Glu Gly Gly
                165                 170                 175
Ser Trp Thr Thr Leu Ala Ser Asp Ile Ser Asp Cys Cys Tyr Leu Thr
            180                 185                 190
Gly Lys Leu Pro Arg Gly Gly Met Tyr Thr Phe Arg Thr Ala Cys Val
        195                 200                 205
Ser Lys Ala Gly Met Gly Pro Tyr Ser Ser Pro Ser Glu Gln Val Leu
    210                 215                 220
```

```
Leu Gly Gly Pro Asn His Leu Ala Ser Glu Glu Glu Ser Ser Arg Gly
225                 230                 235                 240
Arg Pro Ala Gln Leu Leu Pro Ser Thr Lys Thr Phe Ala Phe Gln Thr
                245                 250                 255
Gln Ile Arg Arg Gly Arg Phe Ser Val Val Arg Gln Cys Arg Glu Lys
            260                 265                 270
Ala Ser Gly Arg Ala Leu Ala Ala Lys Ile Val Pro Tyr Gln Pro Glu
        275                 280                 285
Asp Lys Thr Thr Val Leu Arg Glu Tyr Glu Ala Leu Lys Arg Leu His
    290                 295                 300
His Pro His Leu Ala Gln Leu His Ala Ala Tyr Leu Ser Pro Arg His
305                 310                 315                 320
Leu Val Leu Ile Leu Glu Leu Cys Ser Gly Pro Glu Leu Leu Pro Ser
                325                 330                 335
Leu Ala Glu Arg Asp Ser Tyr Ser Glu Ser Asp Val Lys Asp Tyr Leu
            340                 345                 350
Trp Gln Met Leu Ser Ala Thr Gln Tyr Leu His Ala Gln His Ile Leu
        355                 360                 365
His Leu Asp Leu Arg Ser Glu Asn Met Met Val Thr Glu Tyr Asn Leu
    370                 375                 380
Leu Lys Val Ile Asp Leu Gly Asn Ala Gln Ser Leu Ser Gln Glu Lys
385                 390                 395                 400
Val Pro Pro Pro Glu Asn Phe Lys Asp Tyr Leu Glu Thr Met Ala Pro
                405                 410                 415
Glu Leu Leu Glu Gly Gln Gly Ala Val Pro Gln Thr Asp Ile Trp Ala
            420                 425                 430
Ile Gly Val Thr Ala Phe Ile Met Leu Ser Gly Glu Tyr Pro Val Ser
        435                 440                 445
Ser Glu Gly Thr Arg Asp Leu Gln Lys Gly Leu Arg Lys Gly Leu Ile
    450                 455                 460
Gln Leu Ser Arg Cys Tyr Ala Gly Leu Ser Gly Gly Ala Val Ala Phe
465                 470                 475                 480
Leu Gln Ser Ser Leu Cys Ala Arg Pro Trp Gly Arg Pro Cys Ala Ser
                485                 490                 495
Thr Cys Leu Gln Cys Gly Trp Leu Thr Glu Glu Gly Pro Thr Gly Ser
            500                 505                 510
Arg Pro Thr Pro Val Thr Phe Pro Thr Ala Arg Leu Arg Ala Phe Val
        515                 520                 525
Arg Glu Arg Glu Lys Arg Arg Ala Leu Leu Tyr Lys Lys His Asn Leu
    530                 535                 540
Ala Gln Val Arg
545
```

<210> 9
<211> 548
<212> PRT
<213> Mus musculus

<400> 9

```
Met Ala His Ile Ser Arg Ile Leu Lys Gly Arg Pro Glu Gly Pro Glu
 1               5                  10                  15
Arg Glu Gly Pro Pro Arg Lys Lys Ala Gly Leu Ala Ser Phe Arg Leu
```

```
            20                  25                  30
Ser Gly Leu Lys Gly Arg Asp Gln Ala Pro Ser Phe Leu Arg Glu Leu
        35                  40                  45
Ser Asp Glu Ala Val Val Leu Gly Gln Ser Val Thr Leu Ala Cys Gln
    50                  55                  60
Val Leu Ala Gln Pro Thr Ala Gln Ala Thr Trp Ser Lys Asp Gly Val
65                  70                  75                  80
Leu Leu Glu Ser Ser Gly His Leu Leu Ile Ser Ser Thr Leu Lys Asn
                85                  90                  95
Phe Gln Leu Leu Thr Ile Leu Val Val Lys Glu Glu Asp Leu Gly Thr
            100                 105                 110
Tyr Thr Cys Cys Val Ser Asn Pro Leu Gly Thr Ala Val Thr Thr Gly
        115                 120                 125
Val Leu Arg Lys Ala Glu Arg Pro Ser Ser Ser Pro Arg Pro Glu Val
    130                 135                 140
Gly Glu Leu Tyr Lys Asp Ala Val Leu Leu Val Trp Lys Pro Val Glu
145                 150                 155                 160
Ser Cys Gly Pro Val Thr Tyr Ile Val Gln Cys Cys Ile Glu Gly Gly
                165                 170                 175
Ser Trp Thr Thr Leu Ala Ser Asp Ile Ser Asp Cys Cys Tyr Leu Thr
            180                 185                 190
Gly Lys Leu Ser Arg Gly Gly Met Tyr Ile Phe Arg Thr Ala Cys Val
        195                 200                 205
Ser Lys Ala Gly Met Gly Pro Tyr Ser Ser Pro Ser Glu Gln Val Leu
    210                 215                 220
Leu Gly Gly Pro Asn His Leu Ala Ser Glu Glu Glu Ser Ser Arg Gly
225                 230                 235                 240
Arg Pro Ala Gln Leu Leu Pro Ser Thr Lys Thr Phe Ala Phe Gln Met
                245                 250                 255
Gln Ile Arg Arg Gly Arg Phe Ser Val Val Arg Gln Cys Arg Glu Lys
            260                 265                 270
Ala Ser Gly Arg Ala Leu Ala Ala Lys Ile Val Pro Tyr Gln Pro Glu
        275                 280                 285
Asp Lys Thr Ala Val Leu Arg Glu Tyr Glu Ala Leu Lys Arg Leu His
    290                 295                 300
His Pro His Leu Ala Gln Leu His Ala Ala Tyr Leu Ser Pro Arg His
305                 310                 315                 320
Leu Val Leu Ile Leu Glu Leu Cys Ser Gly Pro Glu Leu Leu Pro Ser
                325                 330                 335
Leu Ala Glu Arg Glu Ser Tyr Ser Glu Ser Asp Val Lys Asp Tyr Leu
            340                 345                 350
Trp Gln Met Leu Ser Ala Thr Gln Tyr Leu His Ala Gln His Ile Leu
        355                 360                 365
His Leu Asp Leu Arg Ser Glu Asn Met Met Val Thr Glu Tyr Asn Leu
    370                 375                 380
Leu Lys Val Ile Asp Leu Gly Asn Ala Gln Ser Leu Asp Gln Glu Lys
385                 390                 395                 400
Val Pro Ala Pro Glu Asn Phe Lys Asp Tyr Leu Glu Thr Met Ala Pro
                405                 410                 415
Glu Leu Leu Glu Gly Gln Gly Ala Val Pro Gln Thr Asp Ile Trp Ala
            420                 425                 430
Ile Gly Val Thr Ala Phe Ile Met Leu Ser Gly Glu Tyr Pro Glu Ser
        435                 440                 445
Ser Glu Gly Thr Arg Asp Leu Gln Lys Gly Leu Arg Lys Gly Leu Ile
    450                 455                 460
Arg Leu Ser Arg Cys Tyr Ala Gly Leu Ser Gly Gly Ala Val Ala Phe
465                 470                 475                 480
```

```
Leu Gln Ser Ser Leu Cys Ala Gln Pro Trp Gly Arg Pro Cys Ala Ser
                485                 490                 495
Thr Cys Leu Gln Cys Gly Trp Leu Thr Glu Glu Gly Pro Thr Gly Ser
            500                 505                 510
Arg Pro Thr Pro Val Thr Phe Pro Thr Val Arg Leu Arg Ala Phe Val
        515                 520                 525
Arg Glu Arg Glu Lys Arg Arg Ala Leu Leu Tyr Lys Lys His Asn Leu
    530                 535                 540
Ala Gln Val Arg
545
```

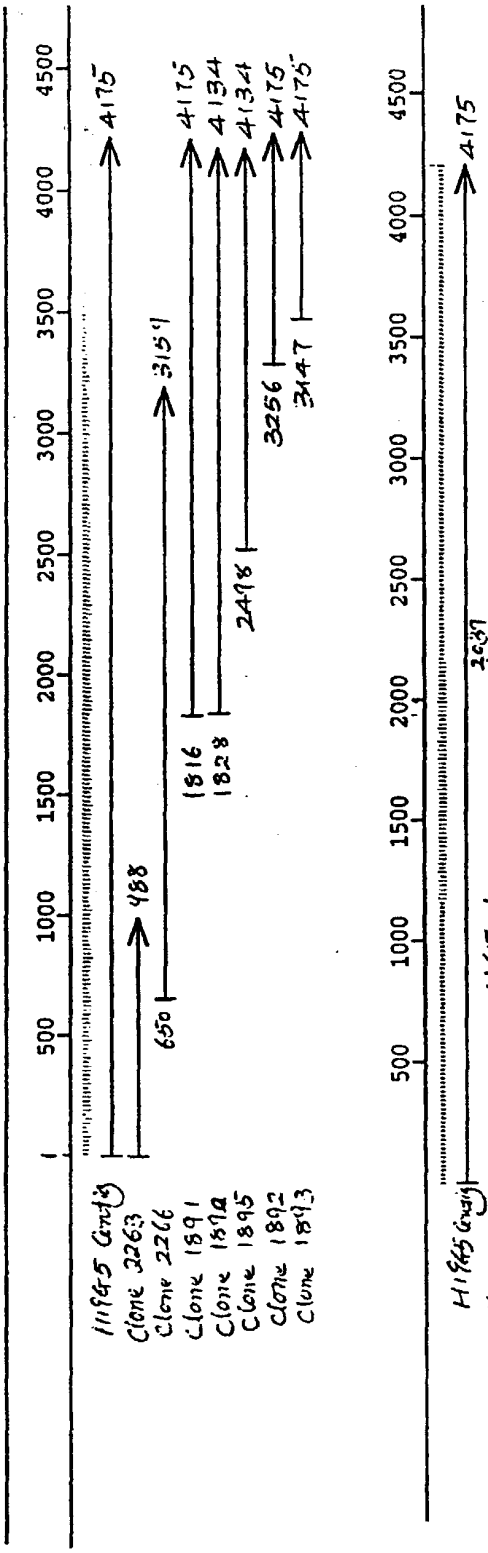

1
500
1000
1500
2000
2500
3000
3500
4000
4500
111845 Contig
Clone 2263
Clone 2266
Clone 1891
Clone 1894
Clone 1895
Clone 1892
Clone 1893
4175
488
650
3151
1816
1828
4175
4134
4134
2498
3256
4175
3147
4175
H19655 Contig
Clone 2264
1167
2689
4175

FIG. 1

Alignment Report of Untitled, using Clustal method with PAM250 residue weight table.
Monday, April 3, 2000 5:11 PM

Page 1

```
      MA--------------------------------------  Majority
               10        20        30        40
1     MGPGDISLPGRPKPGPCSSPGSASQASSSQVSSLRVGSSQ  H19G5-F1.pep
1     MA--------------------------------------  R19G5-S.pep
1     MA--------------------------------------  M19G5-S.pep

      ----------------------------------------  Majority
               50        60        70        80
41    VGTEPGPSLDAEGWTQEAEDLSDSTPTLQRPQEQATMRKF  H19G5-F1.pep
3     ----------------------------------------  R19G5-S.pep
3     ----------------------------------------  M19G5-S.pep

      ----------------------------------------  Majority
               90       100       110       120
81    SLGGRGGYAGVAGYGTFAFGGDAGGMLGQGPMWARIAWAV  H19G5-F1.pep
3     ----------------------------------------  R19G5-S.pep
3     ----------------------------------------  M19G5-S.pep

      ----------------------------------------  Majority
              130       140       150       160
121   SQSEEEEQEEARAESQSEEQQEARAESPLPQVSARPVPEV  H19G5-F1.pep
3     ----------------------------------------  R19G5-S.pep
3     ----------------------------------------  M19G5-S.pep

      ----------------------------------------  Majority
              170       180       190       200
161   GRAPTRSSPEPTPWEDIGQVSLVQIRDLSGDAEAADTISL  H19G5-F1.pep
3     ----------------------------------------  R19G5-S.pep
3     ----------------------------------------  M19G5-S.pep

      ----------------------------------------  Majority
              210       220       230       240
201   DISEVDPAYLNLSDLYDIKYLPFEFMIFRKVPKSAQPEPP  H19G5-F1.pep
3     ----------------------------------------  R19G5-S.pep
3     ----------------------------------------  M19G5-S.pep

      ----------------------------------------  Majority
              250       260       270       280
241   SPMAEEELAEFPEPTWPWPGELGPHAGLEITEESEDVDAL  H19G5-F1.pep
3     ----------------------------------------  R19G5-S.pep
3     ----------------------------------------  M19G5-S.pep

      ----------------------------------------  Majority
              290       300       310       320
281   LAEAAVGRKRKWSSPSRSLFHFPGRHLPLDEPAELGLRER  H19G5-F1.pep
3     ----------------------------------------  R19G5-S.pep
3     ----------------------------------------  M19G5-S.pep
```

FIG. 2

Alignment Report of Untitled, using Clustal method with PAM250 residue weight table.
Monday, April 3, 2000 5:11 PM


```
      ------HISRILKGRPEGPEKEGPPRKKAGLASFRLSGLK Majority
                    330       340       350       360
321 VKASVEHISRILKGRPEGLEKEGPPRKKPGLASFRLSGLK H19G5-F1.pep
3   ------HISRILKGKPEGPEKEGPPRKKAGLASFRLSGLK R19G5-S.pep
3   ------HISRILKGRPEGPEREGPPRKKAGLASFRLSGLK M19G5-S.pep

      GRDQAPSFLRELSDEAVVLGQSVTLACQVLAQPTAQATWS Majority
                    370       380       390       400
361 SWDRAPTFLRELSDETVVLGQSVTLACQVSAQPAAQATWS H19G5-F1.pep
37  GRDQAPSFLRELSDEAVVLGQSVTLACQVLAQPTAQATWS R19G5-S.pep
37  GRDQAPSFLRELSDEAVVLGQSVTLACQVLAQPTAQATWS M19G5-S.pep

      KDGALLESSGHLLISSTLKNFQLLTILVVXEEDLGTYTCC Majority
                    410       420       430       440
401 KDGAPLESSSRVLISATLKNFQLLTILVVVAEDLGVYTCS H19G5-F1.pep
77  KDGALLESSGHLLISSTLKNFQLLTILVVTEEDLGTYTCC R19G5-S.pep
77  KDGVLLESSGHLLISSTLKNFQLLTILVVKEEDLGTYTCC M19G5-S.pep

      VSNPLGTAVTTGVLRKAERPSSSPRPEVGELYXDAVLLVW Majority
                    450       460       470       480
441 VSNALGTVTTTGVLRKAERPSSSPCPDIGEVYADGVLLVW H19G5-F1.pep
117 VSNPLGTAVTTGVLRKAERPSSSPRPEVGELYTDAVLLVW R19G5-S.pep
117 VSNPLGTAVTTGVLRKAERPSSSPRPEVGELYKDAVLLVW M19G5-S.pep

      KPVESYGPVTYIVQCCIEGGSWTTLASDISDCCYLTGKLS Majority
                    490       500       510       520
481 KPVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLS H19G5-F1.pep
157 KPVESYGPVTYIVQCCIEGGSWTTLASDISDCCYLTGKLP R19G5-S.pep
157 KPVESCGPVTYIVQCCIEGGSWTTLASDISDCCYLTGKLS M19G5-S.pep

      RGGMYTFRTACVSKAGMGPYSSPSEQVLLGGPNHLASEEE Majority
                    530       540       550       560
521 RGGTYTFRTACVSKAGMGPYSSPSEQVLLGGPSHLASEEE H19G5-F1.pep
197 RGGMYTFRTACVSKAGMGPYSSPSEQVLLGGPNHLASEEE R19G5-S.pep
197 RGGMYIFRTACVSKAGMGPYSSPSEQVLLGGPNHLASEEE M19G5-S.pep

      SSRGRPAQLLPSTKTFAFQTQIRRGRFSVVRQCREKASGR Majority
                    570       580       590       600
561 S-QGRSAQPLPSTKTFAFQTQIQRGRFSVVRQCWEKASGR H19G5-F1.pep
237 SSRGRPAQLLPSTKTFAFQTQIRRGRFSVVRQCREKASGR R19G5-S.pep
237 SSRGRPAQLLPSTKTFAFQMQIRRGRFSVVRQCREKASGR M19G5-S.pep

      ALAAKIVPYQPEDKTAVLREYEALKRLHHPHLAQLHAAYL Majority
                    610       620       630       640
600 ALAAKIIPYHPKDKTAVLREYEALKGLRHPHLAQLHAAYL H19G5-F1.pep
277 ALAAKIVPYQPEDKTTVLREYEALKRLHHPHLAQLHAAYL R19G5-S.pep
277 ALAAKIVPYQPEDKTAVLREYEALKRLHHPHLAQLHAAYL M19G5-S.pep
```

Alignment Report of 19G5 align, using Clustal method with PAM250 residue weight table. Page 3
Monday, April 3, 2000 5:15 PM

```
    S P R H L V L I L E L C S G P E L L P S L A E R X S Y S E S D V K D Y L W Q M L  Majority
                    650                 660                 670                 680
640 S P R H L V L I L E L C S G P E L L P C L A E R A S Y S E S E V K D Y L W Q M L  H19G5-F1.pep
317 S P R H L V L I L E L C S G P E L L P S L A E R D S Y S E S D V K D Y L W Q M L  R19G5-S.pep
317 S P R H L V L I L E L C S G P E L L P S L A E R E S Y S E S D V K D Y L W Q M L  M19G5-S.pep

    S A T Q Y L H A Q H I L H L D L R S E N M M V T E Y N L L K V I D L G N A Q S L  Majority
                    690                 700                 710                 720
680 S A T Q Y L H N Q H I L H L D L R S E N M I I T E Y N L L K V V D L G N A Q S L  H19G5-F1.pep
357 S A T Q Y L H A Q H I L H L D L R S E N M M V T E Y N L L K V I D L G N A Q S L  R19G5-S.pep
357 S A T Q Y L H A Q H I L H L D L R S E N M M V T E Y N L L K V I D L G N A Q S L  M19G5-S.pep

    S Q E K V P P P E N F K D Y L E T M A P E L L E G Q G A V P Q T D I W A I G V T  Majority
                    730                 740                 750                 760
720 S Q E K V L P S D K F K D Y L E T M A P E L L E G Q G A V P Q T D I W A I G V T  H19G5-F1.pep
397 S Q E K V P P P E N F K D Y L E T M A P E L L E G Q G A V P Q T D I W A I G V T  R19G5-S.pep
397 D Q E K V P A P E N F K D Y L E T M A P E L L E G Q G A V P Q T D I W A I G V T  M19G5-S.pep

    A F I M L S G E Y P V S S E G T R D L Q K G L R K G L I R L S R C Y A G L S G G  Majority
                    770                 780                 790                 800
760 A F I M L S A E Y P V S S E G A R D L Q R G L R K G L V R L S R C Y A G L S G G  H19G5-F1.pep
437 A F I M L S G E Y P V S S E G T R D L Q K G L R K G L I Q L S R C Y A G L S G G  R19G5-S.pep
437 A F I M L S G E Y P E S S E G T R D L Q K G L R K G L I R L S R C Y A G L S G G  M19G5-S.pep

    A V A F L Q S S L C A Q P W G R P C A S T C L Q C G W L T E E G P T G S R P T P  Majority
                    810                 820                 830                 840
800 A V A F L R S T L C A Q P W G R P C A S S C L Q C P W L T E E G P A C S R P A P  H19G5-F1.pep
477 A V A F L Q S S L C A R P W G R P C A S T C L Q C G W L T E E G P T G S R P T P  R19G5-S.pep
477 A V A F L Q S S L C A Q P W G R P C A S T C L Q C G W L T E E G P T G S R P T P  M19G5-S.pep

    V T F P T A R L R A F V R E R E K R R A L L Y K K H N L A Q V R                  Majority
                    850                 860                 870
840 V T F P T A R L R V F V R N R E K R R A L L Y K R H N L A Q V R                  H19G5-F1.pep
517 V T F P T A R L R A F V R E R E K R R A L L Y K K H N L A Q V R                  R19G5-S.pep
517 V T F P T V R L R A F V R E R E K R R A L L Y K K H N L A Q V R                  M19G5-S.pep
```

Decoration 'Decoration #1': Box residues that match the Consensus exactly.

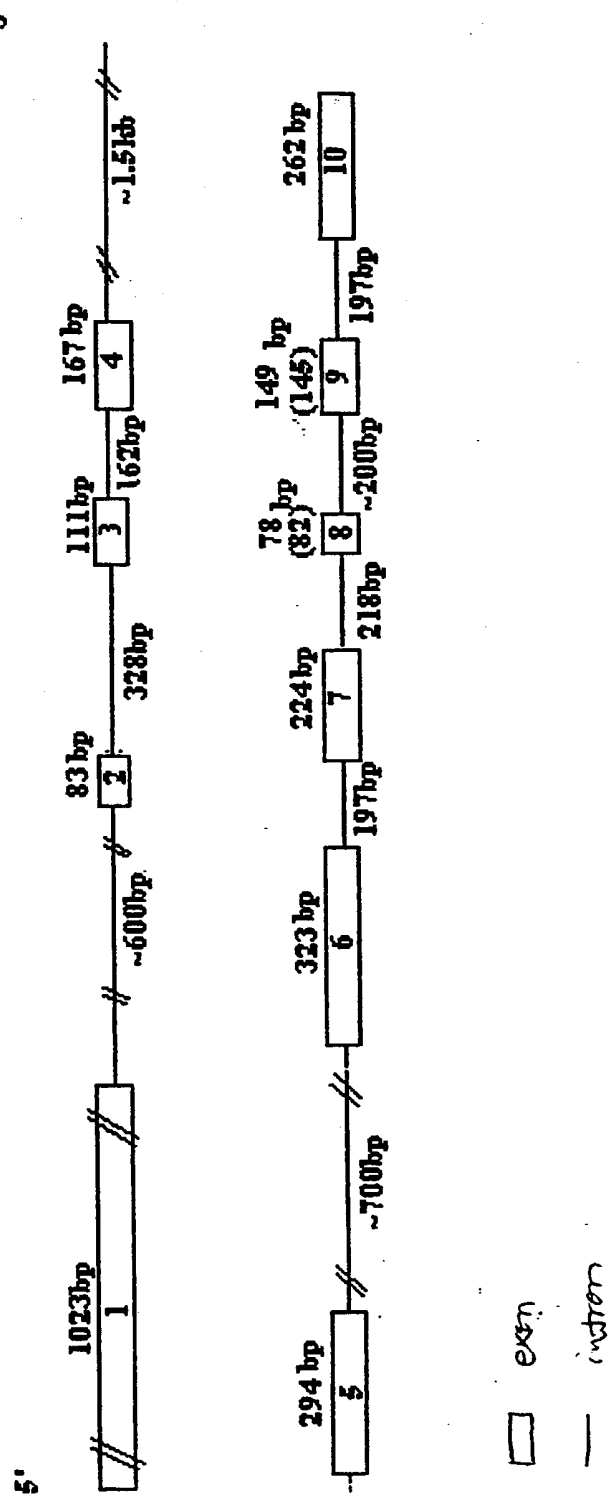

Human H19G5 Gene Structure
5'
3'
1023bp
1
~600bp
83 bp
2
328bp
111bp
3
162bp
167 bp
4
~1.5kb
294 bp
5
~700bp
323 bp
6
197bp
224 bp
7
218bp
78 (82) bp
8
~200bp
149 bp (145)
9
197bp
262 bp
10
exon
intron

FIG. 3

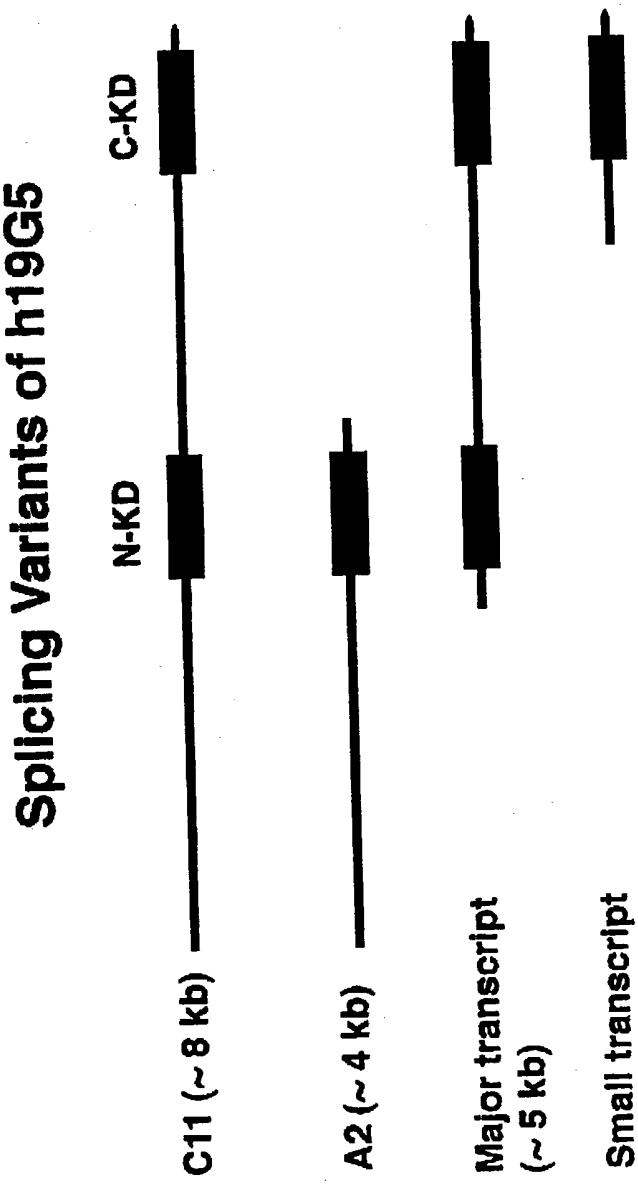

Splicing Variants of h19G5
N-KD
C-KD
C11 (~ 8 kb)
A2 (~ 4 kb)
Major transcript
(~ 5 kb)
Small transcript
(~ 2 kb)

FIG. 4

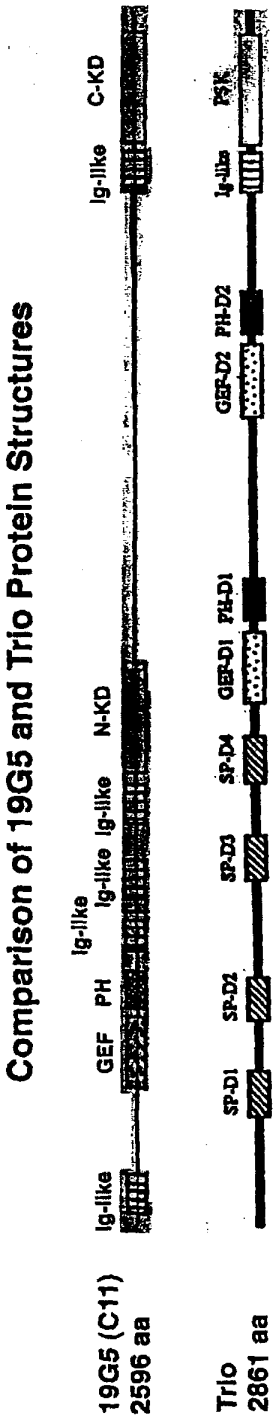

Comparison of 19G5 and Trio Protein Structures
19G5 (C11)
2596 aa
Ig-like
GEF
PH
Ig-like
Ig-like
Ig-like
N-KD
Ig-like
C-KD
Trio
2861 aa
SP-D1
SP-D2
SP-D3
SP-D4
GEF-D1
PH-D1
GEF-D2
PH-D2
Ig-like

FIG. 5

Sequence Comparison of 19G5 , Trio, and MLCK Kinase Domains

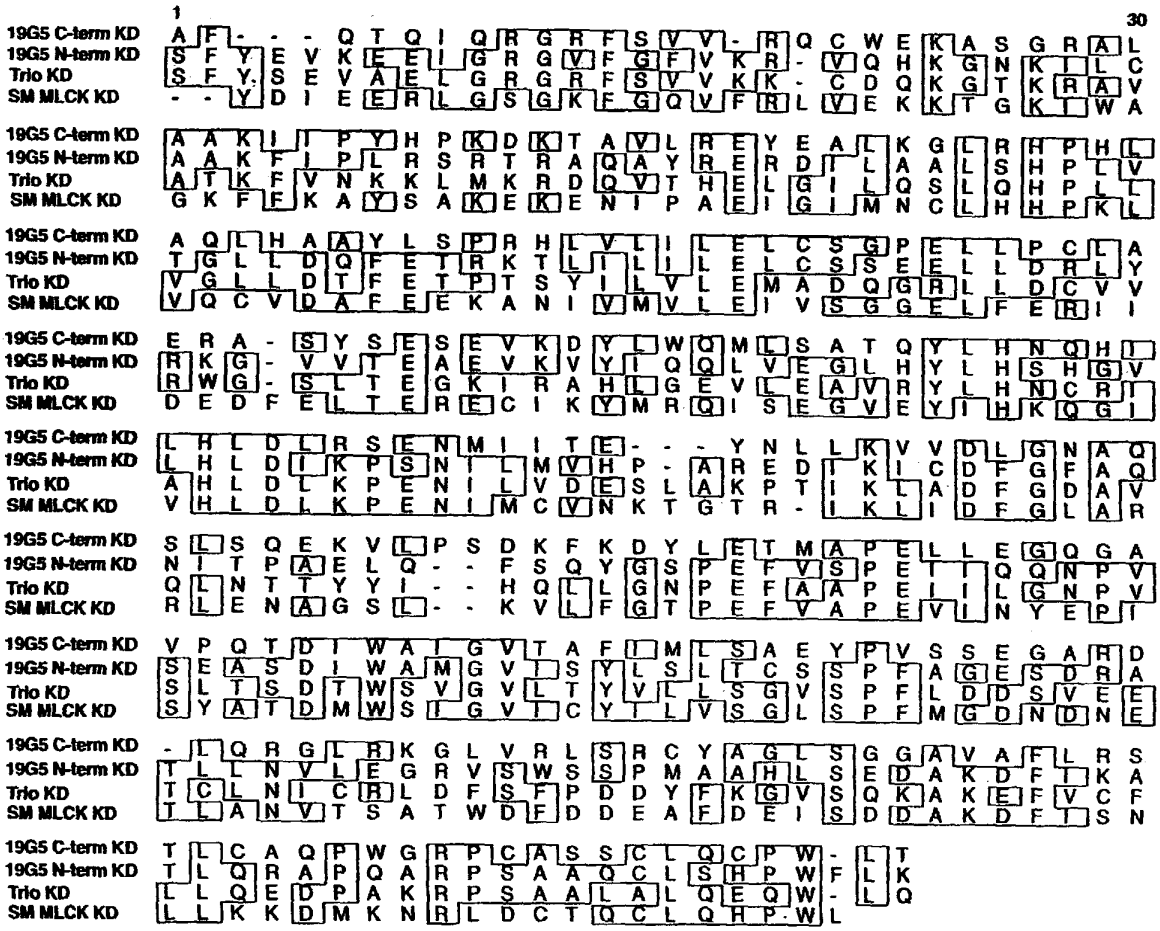


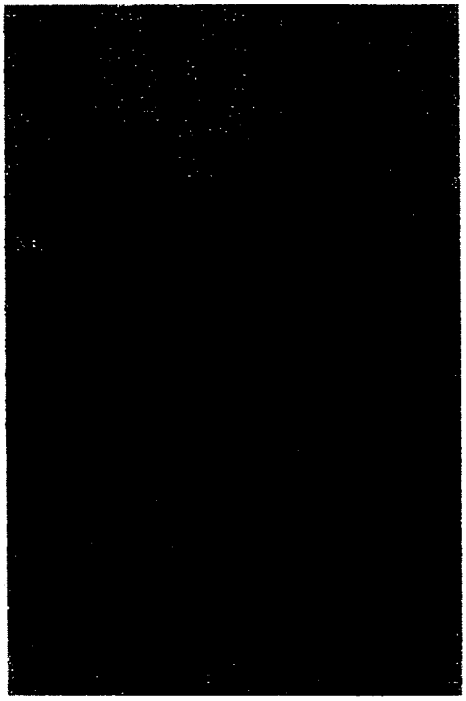
FIG. 6

Sequence Comparison of 19G5 and Trio GEF Domains

FIG. 7

Tissue distribution of Rat H19G5 kinase

Rat MTN Blot Probed with R19G5



Rat MTN Blot Probed with beta-Actin

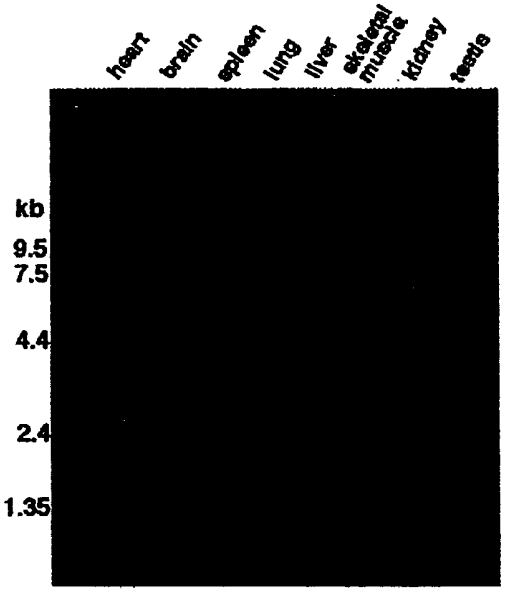


FIG. 8

Human MTN Blot I Probed with H19G5

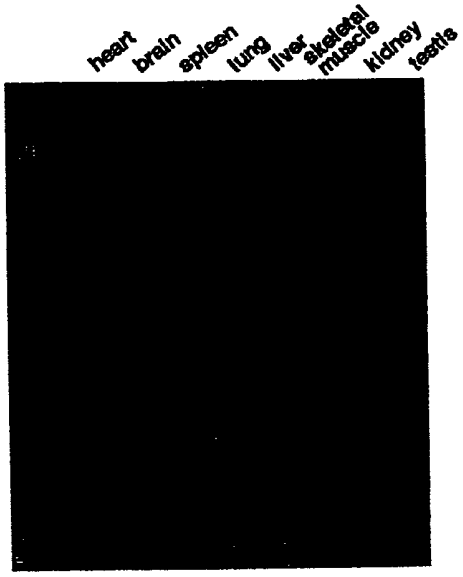


FIG. 9

Genomic Southern Blot with H19G5 3917/3886 fragment as Probe

FIG. 10

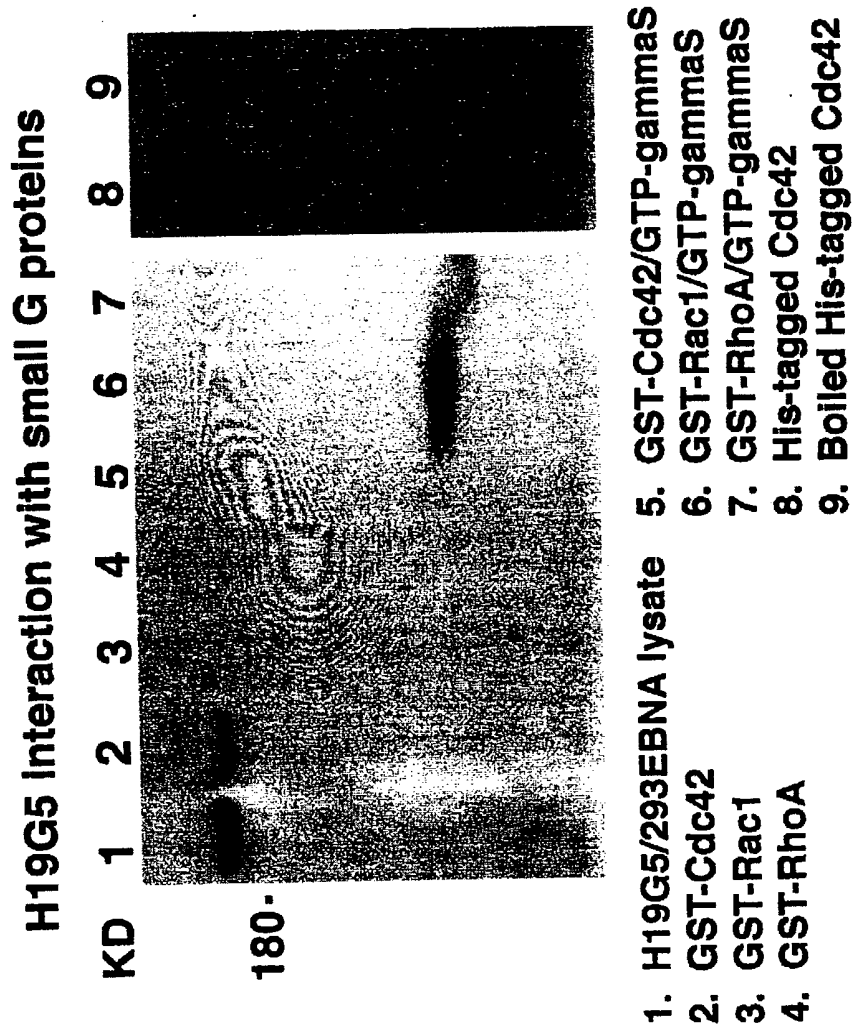

H19G5 Interaction with small G proteins
KD 1 2 3 4 5 6 7 8 9
180-
1. H19G5/293EBNA lysate
2. GST-Cdc42
3. GST-Rac1
4. GST-RhoA
5. GST-Cdc42/GTP-gammaS
6. GST-Rac1/GTP-gammaS
7. GST-RhoA/GTP-gammaS
8. His-tagged Cdc42
9. Boiled His-tagged Cdc42

FIG. 11

19G5-GFP Proteins' Localization In C2C12 Myoblast

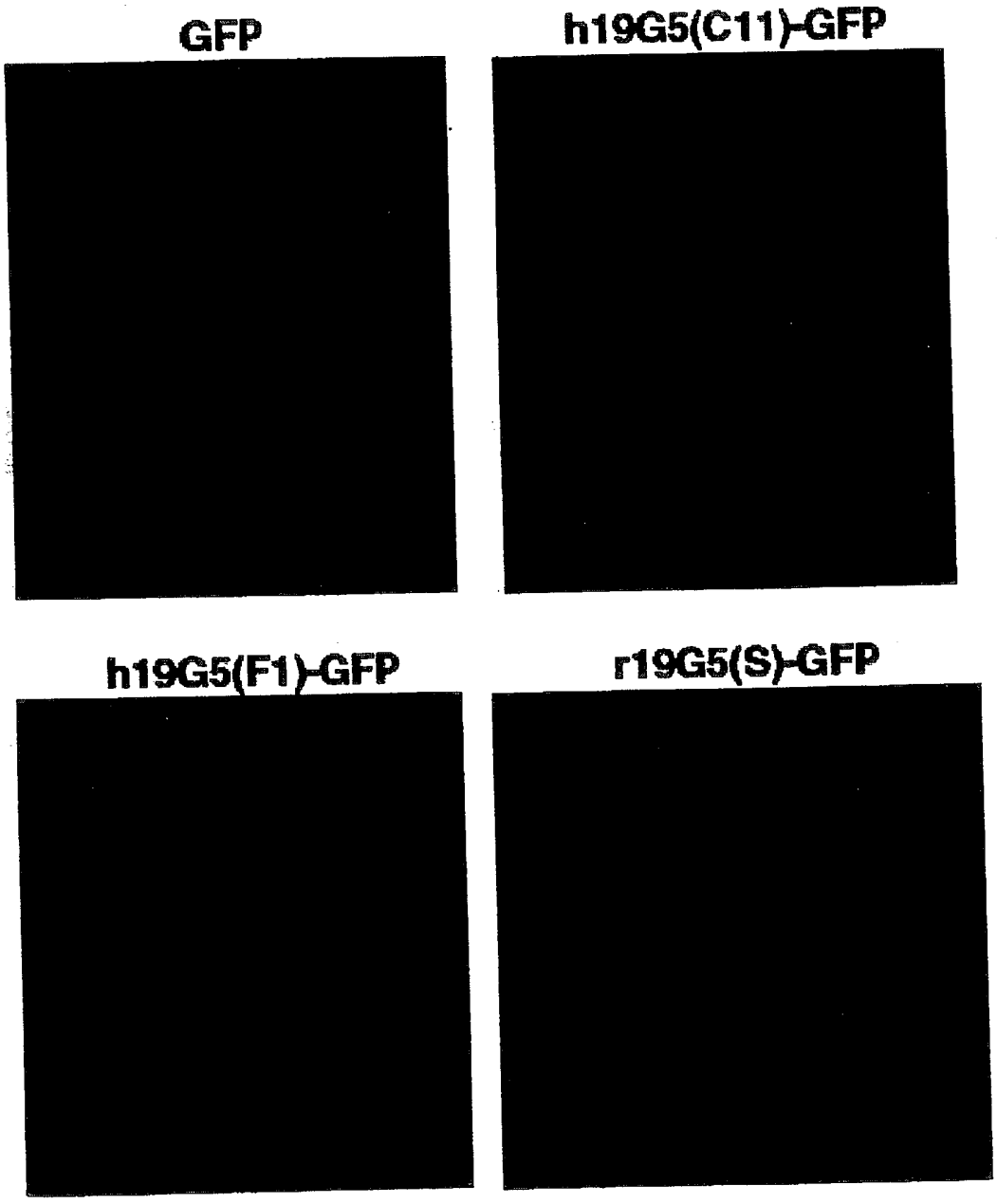


FIG. 12

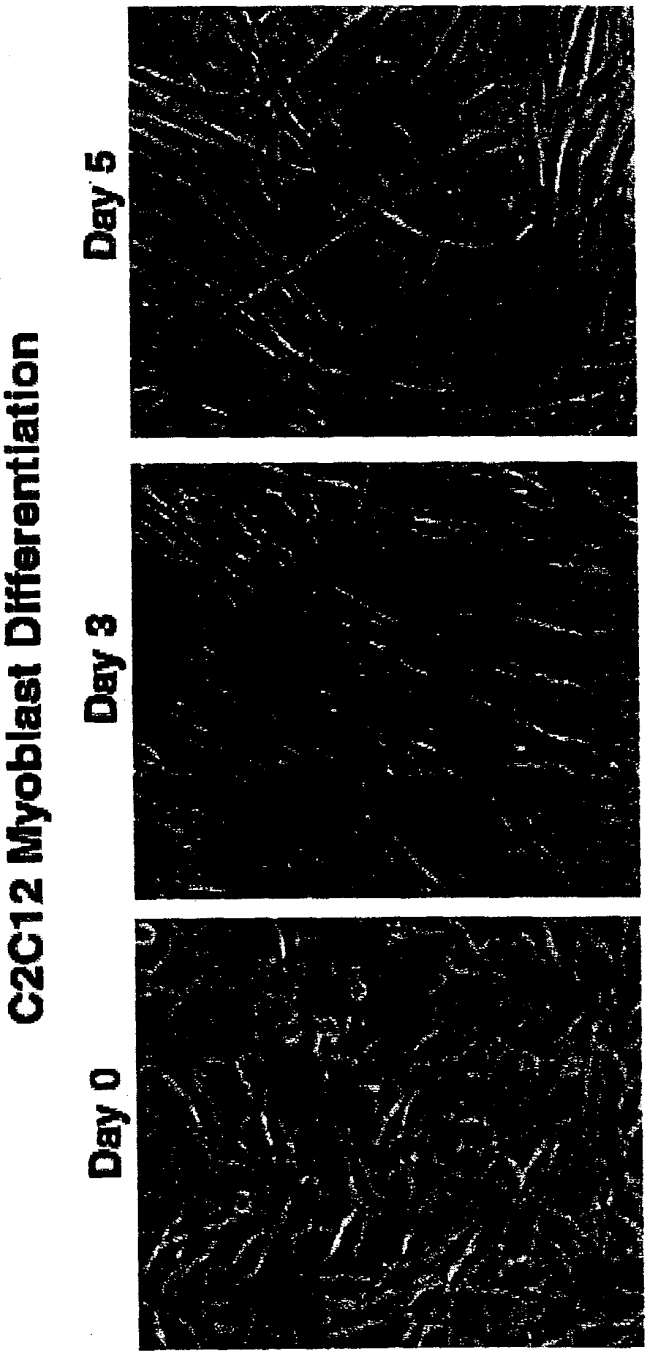


FIG. 13

Myogenin Induction During C2C12 Cell Differentiation

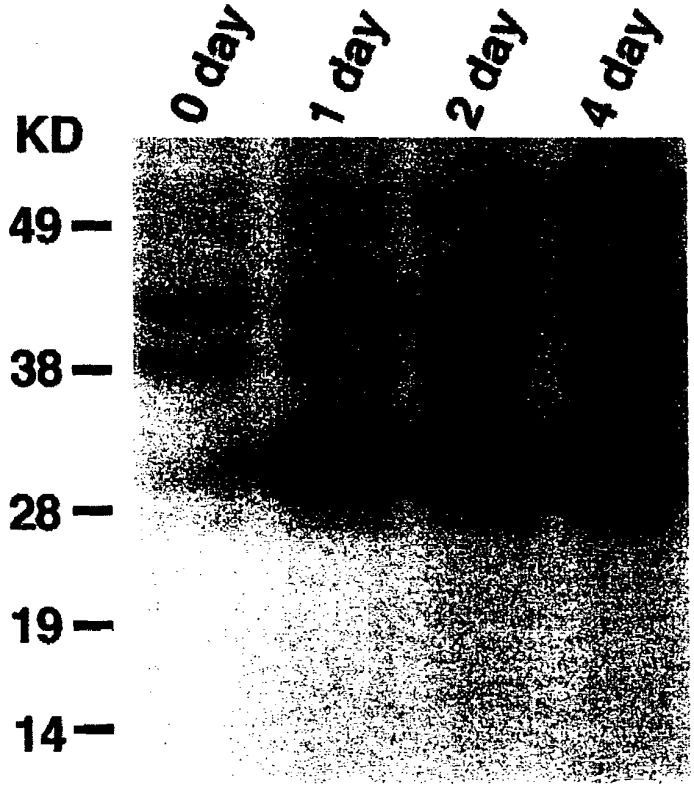


FIG. 14

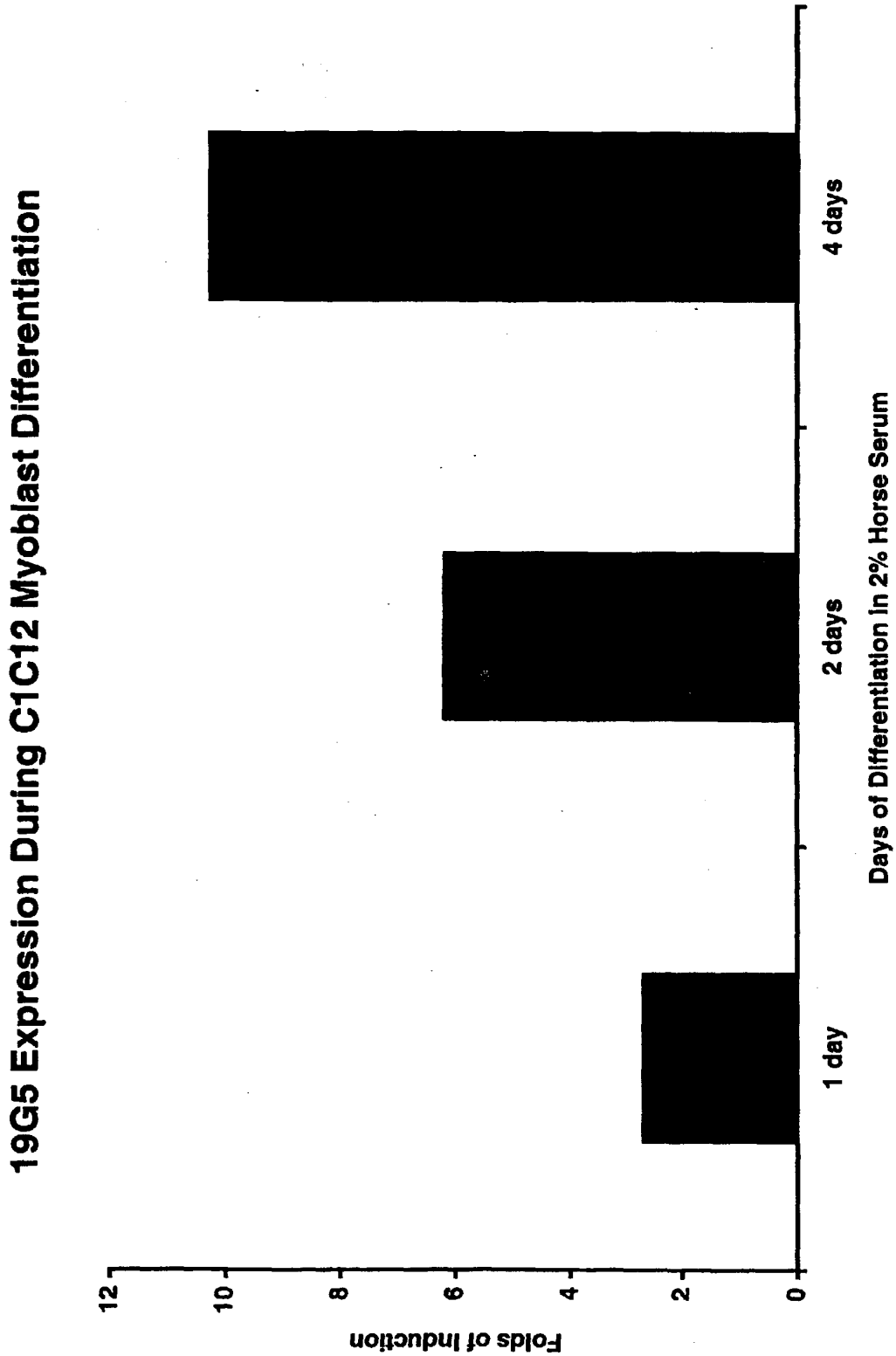

19G5 Expression During C1C12 Myoblast Differentiation
Folds of Induction
0
2
4
6
8
10
12
1 day
2 days
4 days
Days of Differentiation in 2% Horse Serum

FIG. 15

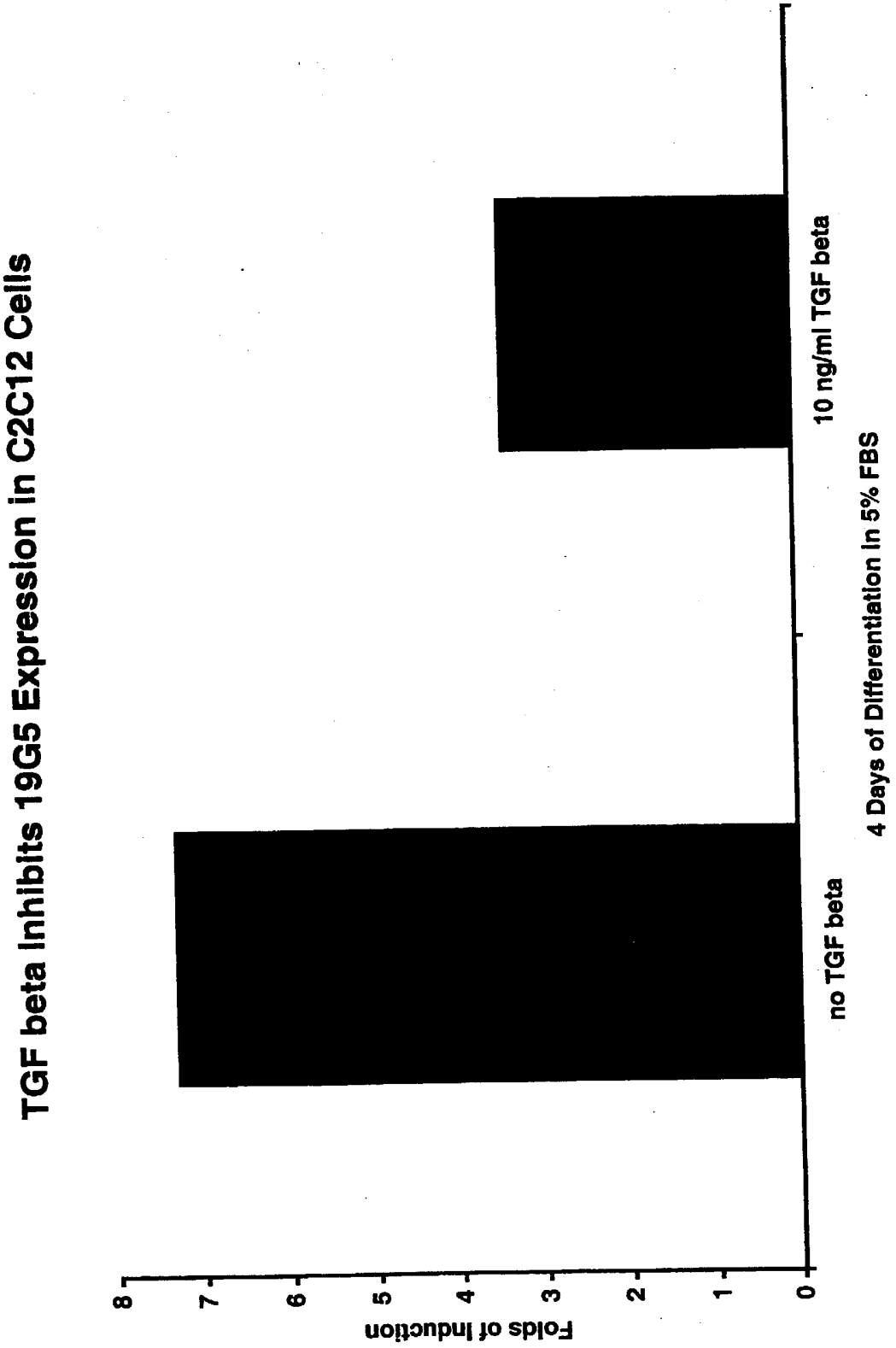

TGF beta Inhibits 19G5 Expression in C2C12 Cells
Folds of Induction
0
1
2
3
4
5
6
7
8
no TGF beta
10 ng/ml TGF beta
4 Days of Differentiation in 5% FBS

FIG. 16

INTERNATIONAL SEARCH REPORT

Int ational Application No
PCT/US 00/09488

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C07K16/18 C12N15/11 G01N33/68
A61K38/17 A61K39/395 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WAYE M. M. Y. ET AL.: "Gene expression of adult human heart as revealed by random sequencing of a cDNA library." EMBL DATABASE ACCESSION NUMBER R45853, 26 May 1995 (1995-05-26), XP002143928 abstract --- "The Genexpress cDNA program." EMBL DATABASE ACCESSION NUMBER Z19064, 15 December 1992 (1992-12-15), XP002143929 abstract --- -/--	13-15 13-15

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 July 2000	10/08/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Mandl, B

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 00/09488

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DEBANT A. ET AL.: "The multidomain protein Trio binds the LAR transmembrane tyrosine phosphatase, contains a protein kinase domain, and has separate rac-specific and rho-specific guanine nucleotide exchange factor domains." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES, vol. 93, no. 11, 1996, pages 5466-5471, XP002143930 1996 ISSN: 0027-8424 the whole document	1-37

Form PCT/ISA/210 (continuation of second sheet) (July 1992)